# Replicator Dynamics of Co–Evolving Networks

**Aram Galstyan** and **Ardeshir Kianercy**
USC Information Sciences Institute
Marina del Rey, California 90292, USA

**Armen Allahverdyan**
Yerevan Physics Institute
Yerevan, Armenia

### Abstract

We propose a simple model of network co–evolution in a game–dynamical system of interacting agents that play repeated games with their neighbors, and adapt their behaviors and network links based on the outcome of those games. The adaptation is achieved through a simple reinforcement learning scheme. We show that the collective evolution of such a system can be described by appropriately defined replicator dynamics equations. In particular, we suggest an appropriate factorization of the agents' strategies that results in a coupled system of equations characterizing the evolution of both strategies and network structure, and illustrate the framework on two simple examples.

## 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 settings, 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 prespecified, often local rules. This approach produced important results on topological features of social, technological and biological networks.

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 likeminded 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 (Gross and Blasius 2008; Goyal 2005; Perc and Szolnoki 2009; Castellano, Fortunato, and Loreto 2009) for a recent survey of the literature).

To describe coupled dynamics of individual attributes and network topology, 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. 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 (Hofbauer and Sigmund 1998; 2003)), replicator equations have been used to model collective learning and adaptation in a systems of interacting self–interested agents (Sato and Crutchfield 2003).

## Background and Related Work

One of the oldest and best studied models of a network is the Erdös–Rényi random graph defined as $G(N; p)$ where $N$ is the number of vertices, and $p$ is the probability of a link between any two vertices. One of the important topological features of graphs is the degree distribution $p_k$, which is the probability that a randomly chosen node has exactly $k$ neighbors. In the large $N$ limit, the degree distribution of the Erdös–Rényi graph is Poissonian, $p_k = e^{-z}z^k/k!$, where $z = pN$ is the average *degree*, or connectivity. While this model adequately describes the percolation transition of the real networks, it fails to account for many properties of real–word networks such as the Internet, social networks or biological networks. In particular, it has been established that many real–world network exhibit what is called a scale–free phenomenon, where the degree distribution follows a power law $p_k \sim k^{-\gamma}$ over a very wide (orders of magnitude) range of $k$.

To account for the statistical deviations of the observed properties of networks from those prescribed by the Erdös–Rényi random graph model, Barabasi and Albert (Barabasi and Albert 1999) proposed a simple model of an evolving network, based on an idea of preferential attachment. In this model, a network grows by addition of new nodes at each time step. Each new node introduced in the system chooses to connect *preferentially* to sites that are already well connected. Thus, nodes that have higher connectivity will add new links with higher rates. It was shown that the network produced by this simple process has an asymptotic scale–free degree distribution of form $p_k \sim k^{-3}$. Recent variations of the initial preferential attachment model include space-inhomogeneous (Bianconi and Barabási 2001) and time-inhomogeneous generalizations of the preferential attachment mechanism (Dorogovtsev and Mendes 2001), ageing and redistribution of the existing links (Dorogovtsev, Mendes, and Samukhin 2000), preferential attachment with memory (Cattuto, Loreto, and Pietronero 2006), evolutionary generalizations of the preferential attachment (Poncela et al. 2008), *etc*.

(Holme and Newman 2006) suggested a model co–evolving networks that combines linking with internal node dynamics. In their model, each node is assumed to hold one of $M$ possible opinions. Initially, the links are randomly distributed among the nodes. Then, at each time step, a randomly chosen node will re–link, with probability $\phi$, one of his links to a node that holds the same opinion. And with probability $1-\phi$, he will change his opinion to agree with the opinion of one of his (randomly chosen) neighbor. Despite the simplicity of those rules, the model was shown to have a very rich dynamical behavior. In particular, while varying the parameter $\phi$, the model undergoes a phase transition from a phase in which opinions are diverse to one in which most individuals hold the same opinion. (Skyrms and Pemantle 2000) suggested a model of adaptive networks where agents are allowed to interact with each other through games, and reinforce links positively if the outcome of the interaction is favorable to them. They showed that even for the simple games, the resulting structural dynamics can be very complex. A review of numerous other models can be found in a recent survey (Castellano, Fortunato, and Loreto 2009).

In addition to abstract statistical models, recent work has addressed the network formation process from the perspective of game–theoretical interactions between self–interested agents (Bala and Goyal 2000; Fabrikant et al. 2003; Anshelevich et al. 2003). In these games each agent tries to maximize her utility consisted of two conflicting preferences – e.g., minimizing the cost incurred by established edges, and minimizing the distance from all the other nodes in the networks. In the simplest scenarios of those games, the topology that corresponds to the Nash equilibrium can be obtained by solving a one–shot optimization problem. In many situations, however, when the actual cost function is more complex, this might not be possible. Furthermore, in realistic situations, agents might have only local information about the network's topology (or utilities of the other agents in the network), so maximizing a global utility is not an option. In this case, the agents can arrive at Nash equilibrium by dynamically adjusting their strategies.

## Dynamics for Co-Evolving Networks

Let us consider a set of agents that play repeated games with each other. Each round of the game proceeds as follows: First, an agent has to choose what other agent he wants to play with. Then, assuming that the other agent has *agreed* to play, the agent has to choose an appropriate action from the pool of available actions. Thus, to define an overall game strategy, we have to specify how an agent chooses a partner for the game and a particular action.

For the sake of simplicity, let us start with three agents, which is the minimum number required for a non–trivial dynamics. Let us differentiate those agents by indices $x$, $y$, and $z$. Here we will focus on the case when the number of actions available to agents is finite. 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 agent adapt to their environment through a simple reinforcement mechanism. Among different reinforcement schemes, here we focus on (stateless) $Q$-learning (Watkins and Dayan 1992). Within this scheme, the agents' strategies are parameterized through so called $Q$–functions that characterize relative utility of a particular strategy. After each round of game, the $Q$ functions are updated according to the following rule:

$$Q^i_{xy}(t+1) = Q^i_{xy}(t) + \alpha[R^i_{x,y} - Q^i_{xy}(t)] \qquad (1)$$

where $R^i_{x,y}$ is the expected reward of agent $x$ for playing action $i$ with agent $y$, and $\alpha$ is a parameter that determines the learning rate (which can be set to $\alpha = 1$ without a loss of generality).

Next, we have to specify how agents choose a particular neighbor and an action based on their $Q$-function. Here we use the Boltzmann exploration mechanism where the probability of a particular choice is given as (Sutton and Barto 1998)

$$p^i_{xy} = \frac{e^{\beta Q^i_{xy}}}{\sum_{\tilde{y},j} e^{\beta Q^j_{x\tilde{y}}}} \qquad (2)$$

Here the inverse *temperature* $\beta = 1/T > 0$ controls exploration/exploitation tradeoff: for $T \to 0$ the agent always choose the action corresponding to the maximum $Q$–value, while for $T \to \infty$ the agents' choices are completely random.

We now assume that the agents interact with each other many times between two consecutive updates of their strategies. In this case, the reward of the $i$–th agent in Equation 1 should be understood in terms of the *average reward*, where the average is taken over the strategies of other agents,

$R^i_{x,y} = \sum_j A^{i,j}_{xy} p^j_{yx}$, where $A^{i,j}_{xy}$ is the reward (payoff) of agent $x$ playing strategy $i$ against the agent $y$ who plays strategy $j$. Note that generally speaking, the payoff might be asymmetric.

We are interested in the continuous approximation to the learning dynamics. Thus, we replace $t+1 \to t+\delta t$, $\alpha \to \alpha\delta t$, and take the limit $\delta t \to 0$ in (1) to obtain

$$\dot{Q}^i_{xy} = \alpha[R^i_{x,y} - Q^i_{xy}(t)] \quad (3)$$

Differentiating 2, using Eqs. 2, 3, and scaling the time $t \to \alpha\beta t$ we obtain the following replicator equation (Sato and Crutchfield 2003):

$$\frac{\dot{p}^i_{xy}}{p^i_{xy}} = \sum_j A^{i,j}_{xy} p^j_{yx} - \sum_{i,j,\tilde{y}} A^{i,j}_{x,\tilde{y}} p^i_{x\tilde{y}} p^j_{\tilde{y}x} + T \sum_{\tilde{y},j} p^j_{x\tilde{y}} \ln \frac{p^j_{x\tilde{y}}}{p^i_{xy}} \quad (4)$$

Equations 4 describes the collective adaptation of the Q–learning agents through repeated game–dynamical interactions. The first two terms indicate that a probability of a playing a particular pure strategy increases with a rate proportional to the overall goodness of that strategy, which mimics fitness-based selection mechanism in population biology (Hofbauer and Sigmund 1998). The second term, which has an entropic meaning, does not have a direct analogue in population biology (Sato and Crutchfield 2003). This term is due to the Boltzmann selection mechanism, and thus, describes the agents' tendency to randomize over their strategies. Note that for $T = 0$ this term disappears and the equations reduce to the conventional replicator system (Hofbauer and Sigmund 1998).

So far, our discussion has been very general. We now make the assumption that the agents strategies can be factorized as follows:

$$p^i_{xy} = c_{xy} p^i_x \,, \ \sum\nolimits_y c_{xy} = 1, \ \sum\nolimits_i p^i_x = 1. \quad (5)$$

Here $c_{xy}$ is the probability that the agent $x$ will initiate a game with the agent $y$, whereas $p^i_x$ is the probability that he will choose action $i$. Thus, the assumption behind this factorization is that the probability that the agent will perform action $i$ does not depend on whom the game is played against.

To proceed further, we substitute 5 in 4, take a summation of both sides in the above equation once over $y$ and then over $i$, and make use of the normalization conditions in Eq. 5 to obtain the following system:

$$\begin{aligned}\frac{\dot{p}^i_x}{p^i_x} &= \sum_{\tilde{y},j} A^{ij}_{x\tilde{y}} c_{x\tilde{y}} c_{\tilde{y}x} p^j_{\tilde{y}} - \sum_{i,j,\tilde{y}} A^{ij}_{x\tilde{y}} c_{x\tilde{y}} c_{\tilde{y}x} p^i_x p^j_{\tilde{y}} \\ &+ T \sum_j p^j_x \ln(p^j_x / p^i_x) \quad (6)\\ \frac{\dot{c}_{xy}}{c_{xy}} &= c_{yx} \sum_{i,j} A^{ij}_{xy} p^i_x p^j_y - \sum_{i,j,\tilde{y}} A^{ij}_{x\tilde{y}} c_{x\tilde{y}} c_{\tilde{y}x} p^i_x p^j_{\tilde{y}} \\ &+ T \sum_{\tilde{y}} c_{x\tilde{y}} \ln(c_{x\tilde{y}} / c_{xy}) \quad (7)\end{aligned}$$

Equations 6 and 7 are the replicator equations that describe the collective and mutual evolution of the agent strategies and the network structure, by taking into account explicit coupling between the strategies and link wights. Our preliminary analysis suggest that this co–evolutionary system can demonstrate a very rich behavior even for simple games. Below we illustrate the framework on two simple examples.

## Examples

Our preliminary results indicate that the co–evolutionary system Equations 6 and 7 can have a very rich behavior even for simple games. The full analysis of those equations will be reported elsewhere. Here we consider two simple examples, and focus on the link dynamics (Eqs. 7), assuming that the agents play Nash equilibrium strategies of the corresponding two–agent game.

Our first example is a simple coordination game with the following (two–agent) payoff matrix:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Thus, agents get a unit reward if they jointly take the first action, and get no reward otherwise.

We assume that the agents always play the same action (e.g., $p^1_x = p^1_y = p^1_z \equiv 1$) yielding a reward $A^{11} = 1$, so we can focus on the link dynamics. Then the equations characterizing the link dynamics are as follows:

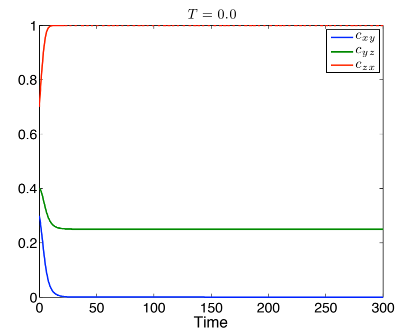
$$\begin{aligned}\frac{\dot{c}_{xy}}{c_{xy}(1-c_{xy})} &= 1 - c_{yz} - c_{zx} + T \ln \frac{1-c_{xy}}{c_{xy}} \quad (8)\\ \frac{\dot{c}_{yz}}{c_{yz}(1-c_{yz})} &= 1 - c_{zx} - c_{xy} + T \ln \frac{1-c_{yz}}{c_{yz}} \quad (9)\\ \frac{\dot{c}_{zx}}{c_{zx}(1-c_{zx})} &= 1 - c_{xy} - c_{yz} + T \ln \frac{1-c_{zx}}{c_{zx}} \quad (10)\end{aligned}$$

We note that the system allows different rest–points some of which correspond to pure Nash equilibria (NE). For instance, one such configuration is $c_{xy} = 1 - c_{yz} = 1$, while $c_{zx}$ can have arbitrary value. In this NE configuration agents $x$ and $y$ always play against each other while agent $z$ is isolated. In addition, there is an interior rest point at $c_{xy} = c_{yz} = c_{zx} = 1/2$, which is again a NE configuration. A simple analysis yields that this symmetric rest point is unstable if the temperature is below a certain critical value. This can be shown by linearizing the system around the symmetric rest point and obtaining the following Jacobian:

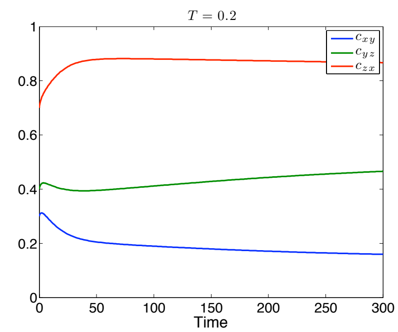
$$J_T = \frac{1}{4} \begin{pmatrix} -4T & -1 & -1 \\ -1 & -4T & -1 \\ -1 & -1 & -4T \end{pmatrix}, \quad (11)$$

It is straightforward to show that for $4T > 1$ all three eigenvalues of this Jacobian become negative, thus making the interior rest point stable. The dynamic of links for various temperature is shown in Figure 1.

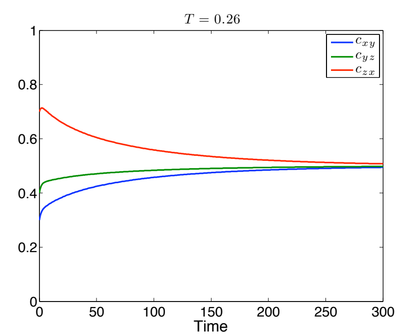
As a second example, we consider the Rock-Paper-Scissor (RPS) normal game which has the following payoff




(a)




(b)




(c)

Figure 1: Dynamics of links for various temperatures.

matrix:

$$A = \begin{pmatrix} \epsilon & -1 & 1 \\ 1 & \epsilon & -1 \\ -1 & 1 & \epsilon \end{pmatrix}$$

where $-1 < \epsilon < 1$. This game has a mixed Nash equilibrium where all the strategies are played with equal probabilities $1/3$. Note that RPS game can have a very rich and interesting dynamics even for two players. For instance, it has been noted that for a two-person RPS game at $T = 0$ the dynamical system might show a chaotic behavior at certain range of $\epsilon$ and never reach an equilibrium (Sato, Akiyama, and Farmer 2002). Again, here we focus on the link dynamics by assuming that the agents play the NE–prescribed strategies $p_x^i = p_y^i = p_z^i = \frac{1}{3}, i = 1, 2, 3$. The equations describing the evolution of the links are as follows:

$$\begin{aligned} \frac{\dot{c}_{xy}}{c_{xy}(1-c_{xy})} &= \frac{\epsilon_x}{3}(1 - c_{yz} - c_{zx}) + T \ln \frac{1-c_{xy}}{c_{xy}} \\ \frac{\dot{c}_{yz}}{c_{yz}(1-c_{yz})} &= \frac{\epsilon_y}{3}(1 - c_{zx} - c_{xy}) + T \ln \frac{1-c_{yz}}{c_{yz}} \\ \frac{\dot{c}_{zx}}{c_{zx}(1-c_{zx})} &= \frac{\epsilon_z}{3}(1 - c_{xy} - c_{yz}) + T \ln \frac{1-c_{zx}}{c_{zx}} \end{aligned}$$

This system has a number of different rest points. For instance, for $-1 < \epsilon < 0$ and $T = 0$, the stable rest point corresponds to a directed triangle with no reciprocated links, $c_{xy} = c_{yz} = c_{zx} = 1$ or $c_{xy} = c_{yz} = c_{zx} = 0$. This is expected, since for $-1 < \epsilon < 0$, the average reward is negative, so the agents are better off not playing with each other at all. There is also an interior rest point at $c_{xy} = c_{yz} = c_{zx} = \frac{1}{2}$. As in the previous example, it can be shown that there is a critical value of $T$ below which this rest point is unstable. Indeed, the Jacobian around the symmetric rest point is as follows:

$$J = \frac{-1}{12} \begin{pmatrix} 12T & \epsilon & \epsilon \\ \epsilon & 12T & \epsilon \\ \epsilon & \epsilon & 12T \end{pmatrix} \qquad (12)$$

A simple calculation shows that the interior rest point becomes stable whenever $T > \epsilon/12$ for $1 > \epsilon > 0$, and $T > |\epsilon|/6$ for $-1 < \epsilon < 0$.

## Discussion

In conclusion, we have presented a replicator–dynamics based framework for studying mutual evolution 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. The examples analyzed here were for simplified scenarios. As a future work, we plan to perform a more through analysis of the dynamics for the fully coupled system. Furthermore, we intend to go beyond the three–agent systems considered here and examine larger systems. Finally, we note that the main premise behind our model is that the strategies can be factorized according to Equations 5. While this assumption is justified for certain type of games, this might not be the case for some other games, where factorization can destroy some NE points that exist in the non–factorized strategy spaces. We intend to examine this question more thoroughly in our future work.

## Acknowledgments

This research was supported by the National Science Foundation under grant No. 0916534.

## References

Anshelevich, E.; Dasgupta, A.; Tardos, E.; and Wexler, T. 2003. Near-optimal network design with selfish agents. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, 511–520. New York, NY, USA: ACM Press.

Bala, V., and Goyal, S. 2000. A noncooperative model of network formation. *Econometrica* 68(5):1181–1229.

Barabasi, A. L., and Albert, R. 1999. Emergence of scaling in random networks. *Science* 286(5439):509–512.

Bianconi, G., and Barabási, A. 2001. Competition and multiscaling in evolving networks. *EPL (Europhysics Letters)* 54:436.

Castellano, C.; Fortunato, S.; and Loreto, V. 2009. Statistical physics of social dynamics. *Reviews of Modern Physics* 81(2):591–646.

Cattuto, C.; Loreto, V.; and Pietronero, L. 2006. Collaborative tagging and semiotic dynamics. *Arxiv preprint cs/0605015*.

Dorogovtsev, S., and Mendes, J. 2001. Language as an evolving word web. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268(1485):2603.

Dorogovtsev, S.; Mendes, J.; and Samukhin, A. 2000. Structure of growing networks with preferential linking. *Physical Review Letters* 85(21):4633–4636.

Fabrikant, A.; Luthra, A.; Maneva, E.; Papadimitriou, C. H.; and Shenker, S. 2003. On a network creation game. In *PODC '03: Proceedings of the twenty-second annual symposium on Principles of distributed computing*, 347–351. New York, NY, USA: ACM Press.

Goyal, S. 2005. Learning in networks. *Group Formation in Economics: Networks, Clubs and Coalitions* 122–70.

Gross, T., and Blasius, B. 2008. Adaptive coevolutionary networks: a review. *Journal of The Royal Society Interface* 5(20):259.

Hofbauer, J., and Sigmund, K. 1998. Evolutionary games and Population Dynamics. *Cambridge University Press*.

Hofbauer, J., and Sigmund, K. 2003. Evolutionary game dynamics. *Journal: Bull. Amer. Math. Soc* 40:479–519.

Holme, P., and Newman, M. E. J. 2006. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* 74(5):056108.

Perc, M., and Szolnoki, A. 2009. Coevolutionary games–A mini review. *BioSystems*.

Poncela, J.; Gómez-Gardeñes, J.; Floría, L.; Sánchez, A.; and Moreno, Y. 2008. Complex cooperative networks from evolutionary preferential attachment. *PLoS One* 3(6).

Sato, Y.; Akiyama, E.; and Farmer, J. 2002. Chaos in learning a simple two-person game. *Proceedings of the National Academy of Sciences of the United States of America* 99(7):4748.

Sato, Y., and Crutchfield, J. 2003. Coupled replicator equations for the dynamics of learning in multiagent systems. *Physical Review E* 67(1):15206.

Skyrms, B., and Pemantle, R. 2000. A dynamic model of social network formation. *Proceedings of the National Academy of Sciences of the United States of America* 97(16):9340–9346.

Sutton, R. S., and Barto, A. G. 1998. *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. The MIT Press.

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