# Adaptive Properties and Memory of a System of Interactive Agents: A Game Theoretic Approach

Roman Gorbunov, Emilia Barakova, Rene Ahn, Matthias Rauterberg
*Designed Intelligence Group, Department of Industrial Design*
*Eindhoven University of Technology*
*Eindhoven, Netherlands*
*r.d.gorbunov@gmail.com*

***Abstract*—In this work we present an evolving system of agents which interact with each other in game theoretic settings. The parameters of the game are considered as time dependent variables representing state of the external environment. The variation of these parameters covers four canonical examples from game theory: prisoners dilemma, hawk-dove, stag-hunt and harmony game. The process of self-adaptation of the proposed system as well as its adaptation to the changing parameters of the environment is considered. We have demonstrated that the introduced system can be in different states which determine the way the system adapts to the external conditions. Stability of these states with respect to the variation of the external conditions has been studied. An ability of the system to accumulate memory about its past experience has been reported.**

***Keywords*-complex adaptive systems; game theory; evolutionary game theory; Turing machines**

## I. Introduction

A complex adaptive system (CAS) is a collection of autonomous, heterogeneous agents, whose behavior is defined by a limited number of rules [1]. In this work we consider a special class of CAS in which agents have a fixed and limited number of actions available to them and outcomes of every pairwise interaction between agents depend on the actions chosen by the two agents participating in the interaction. In more detail, in every pairwise interaction every agent receives a payoff which depends on the actions chosen by the considered agent as well as on the action chosen by another agent involved in the interaction. The mathematical constructs of this kind are typically studied within game theory. However, when the number of agents becomes large, the system starts to exhibit complex and interesting emergent properties which are hard to derive from the underlying simple game theoretic rules of the interactions between the agents and, as a consequence, methods of CAS can be better suited for study of collective properties of the system.

The complexity of the system is conditioned by large number of heterogeneous agents. The behavior of the system as a whole is hard to derive from individual properties of agents and rules of the underlying game. In this framework the payoffs of the game can be considered as parameters of the environment into which the system of interacting agents is embedded. If we supply such a system with a selection mechanism, such that the agents with the largest fitness survive, the populations of the agents of different types will depend on parameters of the game (environment). This property makes the system adaptive. Moreover, the fitness of a given agent in the system depends not only on the parameters of the game but also on the proportion of agents of different types. This makes the system self-adaptive.

| | $A_1$ | $A_2$ |
|---|---|---|
| $A_1$ | $w, w$ | $x, y$ |
| $A_2$ | $y, x$ | $z, z$ |

Table I
Payoff Matrix of a Symmetric Game

In this work we consider a simple game in which every agent has only two actions available. Moreover, we consider symmetric games in which payoffs of two interacting agents depend on their actions and actions of their opponents in the same way. The payoffs in a symmetric game are given by the payoff matrix shown in Table I. The rows and columns of the table correspond to the two actions available to the first and the second agents, respectively. The first and the second number in the cells are payoffs of the first and second agents, respectively. We do not change the logic of a game if we multiply all the payoffs by the same positive number. The logic of the game is also invariant with respect to a constant shift of all payoffs. We also can change the numeration of the actions. Using these three properties we can always set $w = 0$ and $z = 1$. As a consequence we need only two parameters ($x$ and $y$) to completely specify the game.

Changing $x$ and $y$ we will get different games. Four qualitatively different games can be identified. If $x \in (1.0, 2.0)$ and $y \in (0.0, 1.0)$ we have a game which is known in biology and evolutionary game theory as hawk-dove game. In political science and economics the games is more known as chicken game. The earliest presentation of a form of the hawk-dove game was by John Maynard Smith and George Price in their 1973 Nature paper, "The logic of animal conflict" [2]. In biology this game formalizes a situation in which there is a competition for a shared resource and

 

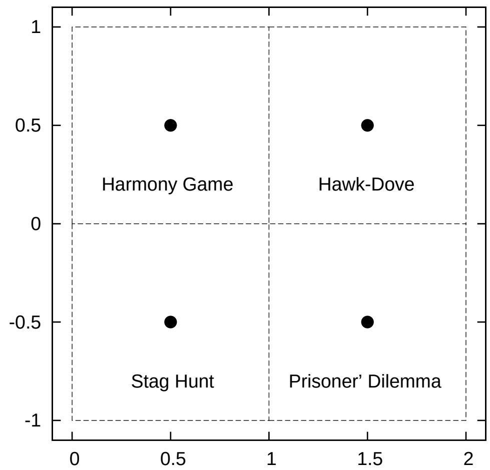


Figure 1. For regions corresponding to four different games

the contestants can choose either conciliation or conflict. Sometimes this game is also referred to as snowdrift game. In game theory this game is known as a canonical model of conflicts between two players and has been the subject of extensive research [3]. The hawk-dove game was also used to compare approaches of CAS and game theory to analysis of systems of interacting agents [1].

If $x \in (1.0, 2.0)$ and $y \in (-1.0, 0.0)$ we get the prisoners dilemma game [4], [5] which is known a canonical tool to study cooperative behavior in game theory, politics, economics, sociology and evolutionary biology. Many natural processes have been abstracted into models in which living beings are engaged in repeated games of prisoner's dilemma. In this game agents need to make a choice between two actions which are called "cooperation" and "defection". The main idea behind the prisoner's dilemma game is that the defection is always more beneficial than the cooperation, independently on what strategy is chosen by the opponent. On the other hand, the mutual cooperation of the agents is more beneficial to both players than the mutual defection. Because of these two properties the game faces players with a dilemma between the cooperation and defection. The iterated prisoner's dilemma has been considered as a part of a model that explains how cooperation can arise in the nature as a result of evolution [6]. The tit-for-tat strategy has been reported as the best deterministic strategy for the iterated prisoner dilemma [7] demonstrating that the prisoner's dilemma can be used to explain reciprocity.

If $x \in (0.0, 1.0)$ and $y \in (-1.0, 0.0)$ we get the stag-hunt game which describes a conflict between safety and social cooperation [8]. This game is also known as "assurance game", "coordination game", and "trust dilemma". Several animal behaviors have been described as stag-hunts including coordination of slime molds and hunting practices of orcas.

If $x \in (0.0, 1.0)$ and $y \in (0.0, 1.0)$ we get the harmony game. This game is trivial since it does not induce any dilemma. It is included into consideration for the completeness of the classification.

In the introduced games the agents always have two choices. These choices can be called as cooperation and defection. By definition the mutual cooperation is always more beneficial than the mutual defection.

## II. METHODS

Additionally to the payoff matrix we need to specify a schema of interaction between agents. There are different ways to organize the interaction between agents. For example, in the so called proposer-responder settings, the first agent, called proposers chooses one of the two actions. The second player, called responder, can see the choice of the proposer and based on that it makes its own choice. After that the choices of the two agents are combined to calculate the payoffs that have to be allocated to the two agents. In our model we take an alternative approach which treats agents in a symmetric way. The two agents make their choices simultaneously. After that their choices are reviled to each other and are used to allocate the payoffs to the agents.

Another important component of the interaction schema is the number of games. If any two agent play with each other only once we have the case of a single stage game or single shot game. Alternatively we could have a case of repeated games. The repeated games can be broadly divided into two classes: finitely and infinitely repeated games. In the case of the finitely repeated games the agents interact with each other a fixed and know number of times. In case of the infinitely repeated games the number of interactions is not fixed. There is a non zero probability, usually fixed and constant, that after the given interaction there will be another interaction. The repeated settings of the game have been shown to be one of the possible mechanisms that can stimulate emergence of cooperative behavior as a result of evolution [9], [10]. In our model we use the settings of the infinitely repeated games. The probability for the next game to happen was set to 0.5.

Additionally to the payoff matrix and the interaction schema we need to specify how agents are paired to interact with each other. One of the options is to assign special location to every agent and let the agent to interact with its neighbors. In this case we get the case of spatial games [11]. The effect of space can have a strong effect on the dynamics of the evolution. In particular agents of certain class can be more successful than other agents because of the fact that they can form spatial clusters in which they mostly interact with the agents of the same kind. This kind of effects supplements the selection of agents by selection of groups of agents and can promote cooperative behavior. In our model we do not consider effects of the spatial locations to focus more on dynamics of relative proportion of different types of agent in the population, which exhibits a complex and interesting behavior worth of separate consideration.

The payoff of an agent depends on what agents were involved into the interactions with the given agent as well as how many games were played with every opponent. In other words the fitness of a given agent is a stochastic property. In the presented model we calculate the average fitness of every agent in the given population of the agents. This assumes that every agent lives long enough to have many rounds of interactions with any other agent. In mathematical terms the fitness of agent $i$ is given by the following equation.

$$f_i = \sum_{j=1}^{n} \nu_j \sum_{k=1}^{\infty} \mu_k \cdot p_{i,j}(k), \quad (1)$$

where the first summation is over all kinds of agents in the population, $\nu_j$ is the portion of agents $j$ in the population. The second sum is over the number of interaction in a single round of interactions between the agents $i$ and $j$. The $\mu_k$ is the probability that there will be $k$ interactions in one round of interactions and $p_{i,j}(k)$ is the payoff of agent $i$ while interacting with agent $j$ in $k$ games.

To start the evolution of the system we need to specify initial relative fractions of agents of different kinds. After that we calculate the average fitness of agents of different kinds using the above given equation (1). Different kinds of agents have different fitness depending on the parameters of the environment (payoff matrix of the game) and portions of the agents of other types. In other words the agents of the same kind can have different fitness in the same game depending on the frequencies of agents of other kind in the population. In this sense the evolutionary process makes the system of agents not only adaptive to the external conditions but also self-adaptive.

On every step of the evolutionary dynamics we calculate average payoff for all kinds of agents in the population. These payoffs determine the changes of the relative proportions of the agents. The proportion of those agents which get a high payoff increases while the proportion of the low-payoff agents decreases. The standard equation giving the changes of the proportions of the agents of different types as a function of their payoffs and proportions of agents of other types is called replicator equation and has the following form:

$$\frac{d\nu_i}{dt} = \nu_i \cdot \left[f_i(\nu_1, \ldots, \nu_n) - f(\nu_1, \ldots, \nu_n)\right], \quad (2)$$

where $f$ is the average payoff in the population:

$$f(\nu_1, \ldots, \nu_n) = \sum_{i=1}^{n} \nu_i \cdot f_i(\nu_1, \ldots, \nu_n) \quad (3)$$

According to this equation the proportion of agents whose payoff is larger than the average one will increase. The growth of a population of agents of a given type is proportional not only on the relative payoff of these agents but also to the current relative size of their population.

The average payoff of the agents grows with the probability of next game in the repeated games settings since the average number of games increases. As a consequence the system will evolve faster if average number of games is larger. To make the average speed of evolution insensitive to the probability of next game we have used a modified version of the replicator equation:

$$\frac{d\nu_i}{dt} = s\nu_i \cdot \frac{f_i(\nu_1, \ldots, \nu_n)}{\max(f_1, \ldots, f_n) - \min(f_1, \ldots, f_n)} \quad (4)$$

According to the modified version of the replicator equation the growth of a subpopulation of agents of a given type depends on the position of their average payoff relative to the maximal and minimal payoffs in the whole population. In front of the right part of the equation we have added a parameter $s$ which can be considered as a time step for the numerical simulation of the evolution. It can be set to a value smaller than 1 to make evolution slower and smoother. In our calculations we use $s = 0.4$.

## III. Results

### A. Evolution with Two Types of Agents

In case of a population consisting of only two types of agents, four qualitatively different kinds of evolutions are possible depending on how payoffs of the agents depend on the portions of the two subpopulations in the whole population. If the average payoff of agent $i$ while interacting with agent $j$ is $p_{ij}$ then payoff of the agent $i$ in the population is equal to:

$$p_i = \nu \cdot p_{i1} + (1 - \nu) \cdot p_{i2}, \quad (5)$$

where $i$ can be 1 or 2 and $\nu$ is the proportion of the first agent in the population. In the Figure 2 we have shown four qualitatively different relations between the payoffs of the agents of the first and second type. In the case of divergence (left top tab) agents of a given type have larger payoffs if their portion is large enough. In this case their portion increases and the agents of another type are completely eliminated from the system. In the case of convergence (right top tab) the agents of a given type have larger payoff than those of the agents of another type if their portion is small enough. As a consequence the portion of those agents which have a larger payoff increases until the convergence points at which both agents get the same payoffs. At this point agents of two types coexist. Another case is absolute domination. This situation happens when one of the two types of agents has larger payoff than another type agents independently on the portion of the agents in the populations. On the left bottom tab we have a case in which agents of the first type absolutely dominate the agents of the second type. In this case the agents of the second type always

 ISBN: 978-1-61208-219-6 

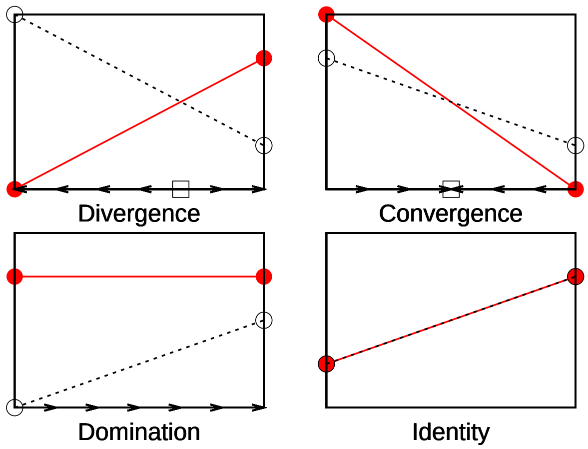


Figure 2. Schematic representation of four possible scenarios of evolution for a system consisting of two kinds of agents. The x-axis is the portion of the agents of the first kind in the population. The red solid and black dashed lines are payoffs of the agents of the first and second kind, respectively.

became extinct. Finally we can have a case when two agents always have the same payoff (see the right bottom tab). In this case the proportions of agents either do not change or exhibit a random walk (depending on the model used for the evolutionary process).

In the prisoners dilemma game defective strategy always dominates the cooperative one since defection is always more beneficial independently on the strategy chosen by the opponent. In contrast, in the harmony game cooperation always dominates defection because the cooperation is more beneficial choice independently on the choice of the opponent. In the hawk-dove game it is beneficial to choose the action which differs from the action chosen by the opponent. This leads to the convergence case. It is more beneficial to be a hawk in a population in which the doves are in the vast majority. If the vast majority of population is hawks, it is more beneficial to be a dove. In the case of the stag-hunt game we have the divergence case. If agents of a certain type are in vast majority, the portion of the agents of another type will decrease and they will become extinct.

### B. Evolution with Eight 1-State Agents

We have considered agents which always choose the same action: defection or cooperation. We can generalize agents if we make their choice dependent on the previous action of the opponent. To define such an agent we need to specify how it should respond on the previous defection or cooperation of its opponent. Moreover, when the agents make the first move in the sequence of the interactions no previous move of the opponent is available. So, we have also to specify what should be the first move of the agent. There are only 8 types of agent of this kind: ccc, ccd, cdc, cdd, dcc, dcd, ddc, and ddd. In our notation the first letter (c or d) denotes the action of the agent in the first move (cooperation and defection, respectively). The second and third letters denote the reaction of the agent on the cooperative and defective moves of the opponent in the previous game, respectively.

For 4 different games we have run evolution more that 8200 times starting from random proportions for the 8 different types of agents. In more detail, for every type of the agent we have generated a random number using a uniform distribution between 0 and 1. These numbers were normalized such that their sum is equal to 1. We used these numbers as initial proportions of the agents of different kinds.

We have found out that in the case of the prisoners dilemma the evolutions converged to one of the two states. The average payoffs of the agents in the two states are 0 and 2. The first state has been observed in 99.94% of cases and the second state has been observed in the rest of the cases. The observed states can also be distinguished by the proportions of the agents of different types.

In the case of the hawk-dove game also two convergence points have been identified. In the first and the second state the average payoff of the agents was equal to 1.15 and 1.93, respectively. The first state has been observed in 88.7% of cases.

In the case of the harmony game only one final state of the evolution has been identified. The average payoff of the agents in this state is equal to 2.

In the case of the stag-hunt game 5 different final states have been identified. The average payoffs of the agents in these states are: 0.00, 0.66, 1.00, 1.33 and 2.00. The frequencies of theses state are 10.47%, 6.70%, 17.66%, 0.35% and 64.82%, respectively.

### C. Dynamics of Average Fitness

It is interesting to consider ability of the system, as a whole, to adapt to the external environment (rules of the game) and extract as much as possible from the environment (in terms of the average payoff of the agents). In other words, the system, as a whole, gets some payoff depending on how the agents (parts of the system) interact with each other and what are the relative portions different types of agents in the system. The system has a mechanism of changing its state (proportion of different kinds of agents) depending on the condition of the external environment. The question that we want to answer is if this mechanism gives the system an ability to adapt positively to the environment and benefit from its conditions.

As a result of the evolution the agents with low fitness become extinct. Does it mean that the average fitness of agents will increase over time as it is prescribed by the Fishers fundamental theorem of natural selection? In many cases the answer is negative. The classical example of the situation when natural selection constantly reduces the average fitness of the population is given by the system in which agents play with each other a single-shot Prisoners dilemma. In this case a population of only cooperators has the highest average fitness, whereas a population of only defectors has the lowest. However, on the individual level, defection is always more beneficial than cooperation and, as a consequence, selections act to increase the relative portion of defectors. After some time, cooperators vanish from the population completely.

This example clearly demonstrate that the system is not always able to adapt positively to the conditions of the environment. However, as we have demonstrated earlier, if we enrich the system by more complex agents, it can exhibit properties of a positive adaptation. For example, in the case of the prisoners dilemma we have found a state in which the average payoff of agents in the system was equal to maximal one. However, this state was very rare. It has been reached only in 0.06% of cases, if we stared the evolution from random populations of agents of different kinds.

### D. Stability of Adaptive Properties

The next question that we wanted to answer is how stable is the state in which the system is able to adapt positively to the priosner's dilemma conditions. To answer this question we started from the higher-payoff state of the prisoners dilemma and have performed a series of jumps to the conditions corresponding to the other types of games. Two sequences of conditions have been considered. The first sequence was: stag-hunt, hawk dove, prisoners dilemma. The second sequence was: hawk dove, stag-hunt, prisoners dilemma. In more detail, first we performed 700 hundreds steps of the evolution under the stag-hunt conditions. This number of steps was sufficient to reach a convergence. After that another 700 hundreds steps of evolution have been performed under the hawk-dove conditions. And finally we performed another 700 steps of evolution in the prisoners dilemma settings to check if the system is able to go back to the original higher-payoff state. In other words, we wanted to find out if evolution under other conditions destroys the ability of the system to adapt positively to the prisoners dilemma settings. To make the test even harder, we have performed a small randomization of the populations between the switching of the conditions of the external environment. The randomization was also performed to include into the evolutionary process under new conditions those types of agents which became completely extinct under the previous phase of the evolution. This test has shown that in 100 out of 100 cases the system was able to keep the ability to adapt positively to the prisoners dilemma. In other words, in spite on the fact that evolution under stag-hunt and hawk-dove conditions as well as the randomization changed the initial proportions of the agents the final evolution of the system, under the prisoners dilemma settings, was always able to converge to the higher-payoff state. Moreover, we found out that the ability to adapt positively to the prisoners dilemma conditions leads to the ability to adapt positively to the stag-hunt and hawk-dove conditions. In the stag-hunt and hawk-dove cases the evolution of the system converged to the highest-payoff state. An example of the positive adaptation is shown in the Figure 3. The red and black dots give an example of a positive and negative adaptation, respectively. The green lines show the allowed average payoffs in the converged state of the corresponding conditions. The test with the changed order of the stag-hunt and hawk-dove games leads us to the same conclusions as in the previous case.

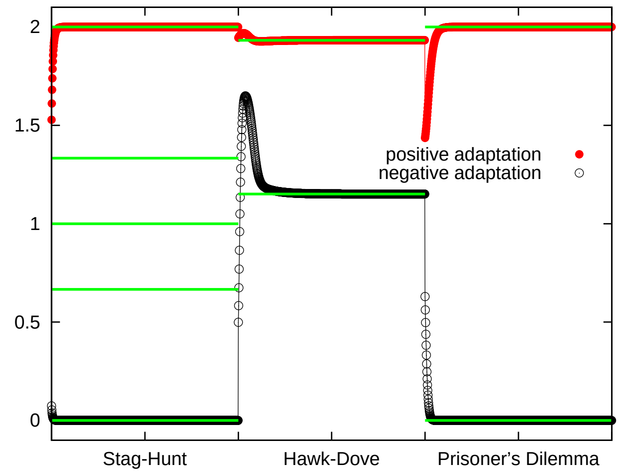


Figure 3. Examples of positive and negative adaptation of the system while changing conditions of the environment

Another interesting property of the considered system is that evolution of the system under new conditions does not completely destroy memory of the system about previous conditions of the evolution. This property can be proven in the following way. If we start from the populations corresponding to the higher-payoff state of the prisoners dilemma and evolve the system in the stag-hunt conditions until the convergence we will end up in the highest-payoff state. If we then evolve this system in the prisoners dilemma settings we will always converge back to the original higher-payoff state of the prisoners dilemma. We have observed this behavior in 100 cases out of 100. In contrast, if start from the random populations and evolve the system in the stag-hunt conditions in some cases we will converge to the higher-payoff state. Then, if we start from the higher-payoff state obtained in this way and start the evolution in the prisoners dilemma settings, we will converge to the higher-payoff state of the prisoners dilemma only in about half of the cases (48 cases out of 100 in our experiment). Thus we can say that the probability that evolution under the prisoners dilemma conditions will converge to the higher-payoff state depends not only on from what state we started the evolution (in our case the higher-payoff state of the stag-hunt conditions) but also on how the initial state was obtained. Despite the fact that in both cases we started the evolution from the higher-payoff state of the stag-hunt conditions, the system had much

 ISBN: 978-1-61208-219-6 

more larger chances to converge to the higher-payoff state of the prisoners dilemma, if the system already was in the higher-payoff state of the prisoners dilemma game. This kind of memory is explained by the fact that a fixed state does not necessarily mean fixed populations of the agents of different types. In other words different populations of the agents of different kinds can correspond to the same state. For example, in the above considered example of the memory we had the higher-payoff state for the stag-hunt conditions. In this case only ccc and ccd agents are present in the system. However, the relative proportion of these two types of agents is not fixed. This degree of freedom can be used to encode the memory of the system.

## IV. Conclusion

In this work we have presented a model of a complex adaptive system based on concepts of game theory. An evolving system of heterogeneous agents interacting with each other in a game theoretic way has bas introduced. The influence of the external environment has been modeled by payoff matrix of the game. The state of the system is given by the proportion of agents of different types in the population. The adaptive properties of the system have been introduced through an evolutionary selection of the agents. An ability of the system to increase an average payoff under given external conditions has been studied as an adaptive process. In this context the importance of the initial state of the system has been demonstrated. The stability of the positive adaptation has been considered. Moreover, it has been demonstrated that the introduced system could have a memory about the history of the changes of the external conditions. It has been shown that some memory about old conditions remains even if the system evolved in new conditions long enough to reach a convergence.

## Acknowledgment

The research reported in this paper is supported by NWO (De Nederlandse Organisatie voor Wetenschappelijk Onderzoek) User Support Program Space Research. The project number is ALW-GO-MG/07-13.

## References

[1] M. Hadzikadic, T. Carmichael, and C. Curtin, "Complex adaptive systems and game theory: An unlikely union," *Complexity*, vol. 16, pp. 34–42, 2010.

[2] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, pp. 15–18, 1973.

[3] A. Rapoport and A. M. Chammah, "The game of chicken," *American Behavioral Scientist*, vol. 10, pp. 10–28, 1966.

[4] M. M. Flood, "Some experimental games," *Management Science*, vol. 5, pp. 5–26, 1958.

[5] M. Dresher, *The Mathematics of Games of Strategy: Theory and Applications*. Prentice-Hall, Englewood Cliffs, NJ., 1961.

[6] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, pp. 1390–1396, 1981.

[7] A. Rapoport and A. M. Chammah, *Prisoner's Dilemma*. University of Michigan Press, 1965.

[8] B. Skyrms, *The Stag Hunt and Evolution of Social Structure*. Cambridge: Cambridge University Press., 2004.

[9] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, pp. 1560–1563, 2006.

[10] R. J. Aumann, "Acceptable points in general cooperative n-person games," *Annals of Mathematics Studies*, vol. 40, pp. 287–324, 1959.

[11] M. A. Nowak, S. Bonhoeffer, and R. M. May, "Spatial games and the maintenance of cooperation," *Proceedings of the National Academy of Sciences*, vol. 91, pp. 4877–4881, 1994.

 ISBN: 978-1-61208-219-6 