The El Farol Bar Problem as an Iterated *N*-Person Game

Miklos N. Szilagyi

Department of Electrical and Computer Engineering
University of Arizona

The El Farol Bar problem is treated as an iterated *N*-person battle of the sexes or leader game for any initial attendance ratio and any value of the bar capacity. It is shown that the bar attendance always fluctuates around the value of the bar capacity. The fluctuations' amplitude is inversely proportional to the total number of the participating agents.

1. Introduction

The famous El Farol Bar problem [1] is an excellent demonstration of the self-referential expectation formation. As such, it serves also as a simple model of financial markets.

A number of agents must decide independently each week whether to go to their favorite bar or not. If an agent expects that the bar will be overcrowded, he will stay at home; otherwise, he will go to the bar. If all agents believe few will go, then all will go, destroying all the expectations. On the other hand, if all agents believe many will go, then nobody will go, destroying the expectations again.

Arthur [1] has shown by using computer simulation based on a sophisticated reasoning that the bar attendance fluctuates rather wildly around the capacity level of the bar. This result was arrived at by many other researchers as well (e.g., [2–4]).

The El Farol problem has been extended to the so-called minority game [5]. The players must choose between two options and those on the minority side win. Many variants of this game have been developed, a large number of papers have been published (e.g., [6–9]), and even books have been written about it [10–12].

In this paper, a much simpler approach is proposed.

2. El Farol as an *N*-Person Battle of the Sexes or Leader Game

We will consider an *N*-person game where the agents have two choices: to go to the bar (*C*) or remain at home (*D*). The bar-goers are rewarded when there are few of them and punished if there are many of them. Those who choose to stay at home are rewarded when the

https://doi.org/10.25088/ComplexSystems.21.2.153

bar is overcrowded and punished if there are few bar-goers because they failed to take advantage of this situation. The reward/penalty (payoff) functions are shown in Figures 1 and 2.

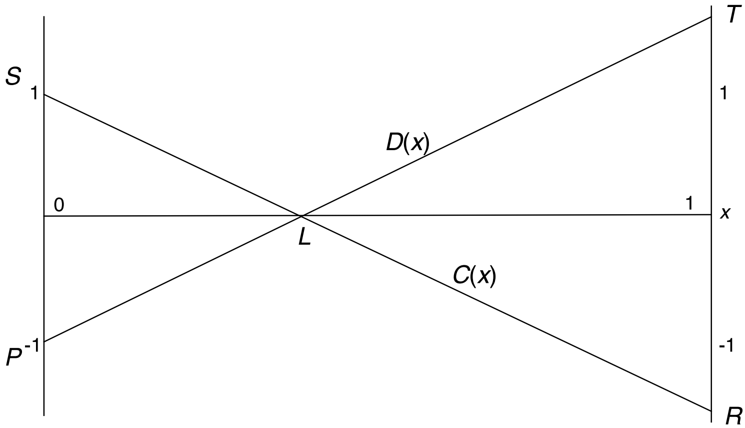

Figure 1. Payoff (reward/penalty) functions for bar-goers (C) and those who stay at home (D) when the bar capacity $L < 0.5$ ($T > S > P > R$). The horizontal axis (x) represents the ratio of the number of bar-goers to the total number of agents; the vertical axis is the reward/penalty.

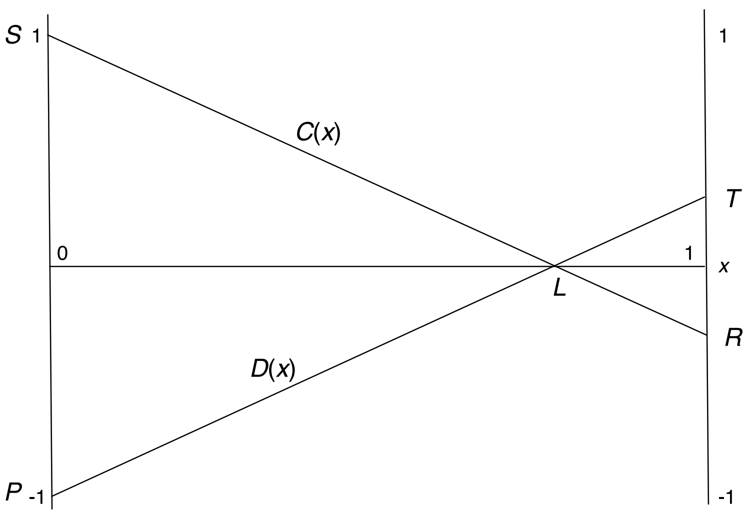

Figure 2. Payoff (reward/penalty) functions for bar-goers (C) and those who stay at home (D) when the bar capacity $L > 0.5$ ($S > T > R > P$). The horizontal axis (x) represents the ratio of the number of bar-goers to the total number of agents; the vertical axis is the reward/penalty.

https://doi.org/10.25088/ComplexSystems.21.2.153

The horizontal axis x in these figures represents the number of bar-goers related to the total number of agents. We will assume that the payoffs $C(x)$ and $D(x)$ are linear functions of this ratio for both choices and the game is uniform, that is, the payoff functions are the same for all agents.

Point $P = D(0)$ corresponds to the payoff when all agents stay at home, $R = C(1)$ is the payoff when all agents go to the bar, S is the reward for going to the bar when everyone else stays at home, and T is the reward for staying at home when everybody else goes to the bar. $C(0)$ and $D(1)$ are impossible by definition, but we will follow the generally accepted notation by extending both lines to the full range of $0 \leq x \leq 1$ and denoting $C(0) = S$ and $D(1) = T$ that makes it simpler to define the payoff functions. For a large number of agents, this extension is not even noticeable.

We arbitrarily choose $S = 1$ and $P = -1$, then connect by straight lines point S with point R (bar-goers' payoff function C) and point P with point T (home-sitters' payoff function D). The bar capacity L is a variable at which value reward becomes punishment and vice versa. Thus the payoff to each agent depends on its choice, on the distribution of other players among bar-goers and home-sitters, and on the value of L.

If $L < 0.5$, then the absolute values of both R and T are greater than one (Figure 1). In this case, $T > S > P > R$, which corresponds to the battle of the sexes game [13]. When $L > 0.5$, then the absolute values of both R and T are less than one (Figure 2). In this case, $S > T > R > P$, which is a leader game [14].

3. Agent-Based Simulation

We will use our agent-based simulation tool developed for social and economic experiments with any number of decision makers operating in a stochastic environment [15] for the investigation of this problem.

Our model has three distinctive features:

1. It is a genuine multi-agent model. It is not based on repeated two-person games.
2. It is a general framework for inquiry in which the properties of the environment as well as those of the agents are user-defined parameters and the number of interacting agents is theoretically unlimited. This model is well suited for simulating the behavior of artificial societies of large numbers of agents.
3. Biological objects including human beings are not always rational. Human behavior can be best described as stochastic but influenced by personality characteristics. In view of this hypothesis, it becomes crucially important to investigate the role of personalities in games. Our agents have various distinct, user-defined personalities.

https://doi.org/10.25088/ComplexSystems.21.2.153

The participating agents are described as stochastic learning cellular automata, that is, as combinations of cellular automata [16, 17] and stochastic learning automata [18, 19]. The cellular automaton format describes the environment in which the agents interact. In our model, this environment is not limited to the agents' immediate neighbors: the agents may interact with all other agents simultaneously. Stochastic learning rules provide more powerful and realistic results than the deterministic rules usually used in cellular automata. Stochastic learning means that behavior is not determined but only shaped by its consequences, that is, an action of the agent will be more probable but still not certain after a favorable response from the environment.

The model and its implementation is described in detail in [15]. We will only briefly explain its most important features here. The software package is available from the author upon request.

A realistic simulation model of a multi-person game must include a number of parameters that define the game to be simulated. Our model has the following user-defined parameters:

1. Size and shape of the simulation environment (array of agents)
2. Definition of neighborhood: the number of layers of agents around each agent that are considered its neighbors
3. Payoff (reward/penalty) functions
4. Updating schemes (learning rules) for the agents' subsequent actions
5. Personalities
6. Initial probabilities of choosing C
7. Initial actions of the agents

Our simulation environment is a two-dimensional array of the participating agents. Its size is limited only by the computer's virtual memory. The behavior of a few million interacting agents can easily be observed on the computer's screen.

There are two actions available to each agent. Each agent must repeatedly choose between them. To be consistent with the terminology accepted in the game theory literature, we call these actions cooperation and defection. Each agent has a probability distribution for the two possible actions. The agents as stochastic learning cellular automata take actions according to their probabilities updated on the basis of the reward/penalty received from the environment for their previous actions, their neighbors' actions, and the agents' personalities. The updating occurs simultaneously for all agents. In the present case, the bar-goers are the cooperators and the home-sitters are the defectors.

The updated probabilities lead to new decisions by the agents that are rewarded/penalized by the environment. After each iteration, the software tool draws the array of agents in a window on the comput-

https://doi.org/10.25088/ComplexSystems.21.2.153

er's screen, with each agent in the array colored according to its most recent action. The experimenter can view and record the evolution of the society of agents as it changes with time.

The updating scheme is different for different agents. Agents with completely different personalities can be allowed to interact with each other in the same experiment. Agents with various personalities and various initial states and actions can be placed anywhere in the two-dimensional array.

A variety of personality profiles and their arbitrary combinations can be represented in the model, including the following:

1. *Pavlovian*. The probability of cooperation p changes by an amount proportional to its reward or penalty from the environment for its previous action (the coefficient of proportionality is the learning rate that is a user-defined parameter).
2. *Stochastically predictable*. p is a constant. For example,
   (a) *Angry*. Always defects ($p = 0$).
   (b) *Benevolent*. Always cooperates ($p = 1$).
   (c) *Unpredictable*. Acts randomly ($p = 0.5$).
3. *Accountant*. p depends on the average reward for previous actions.
4. *Conformist*. Imitates the action of the majority.
5. *Greedy*. Imitates the neighbor with the highest reward.

Aggregate cooperation proportions are changing in time, that is, over subsequent iterations. The iterations are repeated until a certain pattern appears to remain stable or oscillating.

The payoff (reward/penalty) functions are given as two curves: one (C) for a cooperator and another (D) for a defector. The payoff to each agent depends on its choice, on the distribution of other players among cooperators and defectors, and also on the properties of the environment. The payoff curves are functions of the ratio x of cooperators to the total number of neighbors:

$$C(x) = a_c x^2 + b_c x + c_c + d_c \,\text{rnd} \qquad \text{for cooperators} \tag{1}$$

and

$$D(x) = a_d x^2 + b_d x + c_d + d_d \,\text{rnd} \qquad \text{for defectors} \tag{2}$$

where the choice of the coefficients determine the payoff functions. Stochastic factors d_c and d_d can be specified to simulate stochastic responses from the environment; rnd is a random number between 0 and 1. Thus the fourth terms of equations (1) and (2) determine the thickness of the payoff functions. In this simulation, we chose zero stochastic factors, that is, a deterministic environment. The freedom of using quadratic functions for the determination of the reward/penalty system makes it possible to simulate a wide range of social situations, including those where the two curves intersect each other as in the present case (see Figures 1 and 2).

The agents take actions according to probabilities updated on the basis of the reward/penalty received for their previous actions and of their personalities. The updating scheme may be different for different agents. This means that agents with different personalities can be allowed to interact with each other in the same experiment. Agents with various personalities and various initial states and actions can be placed anywhere in the array. The response of the environment is influenced by the actions of all participating agents.

The number of neighborhood layers around each agent and the agent's location determine the number of its neighbors. The depth of agent *A*'s neighborhood is defined as the maximum distance, in the horizontal and vertical directions, that agent *B* can be from agent *A* and still be in its neighborhood. We do not wrap around the boundaries; therefore, an agent in the corner of the array has fewer neighbors than one in the middle. The neighborhood may extend to the entire array of agents.

It is also very important to properly describe the environment. In our model, even in the almost trivial case when both payoff curves are straight lines and the stochastic factors are both zero, four parameters specify the environment. Attempts to describe it with a single variable [20, 21] are certainly too simplistic.

4. Pavlovian Agents

Most biological objects modify their behavior according to Pavlov's experiments and Thorndike's law of conditioning [22]: if an action is followed by a satisfactory state of affairs, then the tendency to produce that particular action is reinforced. These agents are primitive enough not to know anything about their rational choices. However, they have enough "intelligence" to follow Thorndike's law. Their probability of cooperation changes by an amount proportional to the reward/penalty received from the environment. Kraines and Kraines [23], Macy [24], Flache and Hegselmann [25], and others used Pavlovian agents for the investigation of iterated two-person games.

A linear updating scheme is used for the Pavlovian agents: the change in the probability of choosing the previously chosen action again is proportional to the reward/penalty received from the environment (payoff curves):

$$P_{\text{new}} = P_{\text{previous}} + a * \text{reward/penalty} \tag{3}$$

where P_{previous} is the probability of the previous action, P_{new} is the probability of choosing the previously chosen action again, a is the user-defined learning factor, and reward/penalty is $C(x)$ if the previ-

https://doi.org/10.25088/ComplexSystems.21.2.153

ous action was cooperation and $D(x)$ if it was defection. Evidently, a high probability does not guarantee a certain action. Of course, the probabilities always remain in the interval between 0 and 1.

Assume that in a society of Pavlovian agents, the ratio of cooperators is x. Accordingly, the ratio of defectors is $(1-x)$. They are distributed randomly over the two-dimensional array at a certain time. Then $x\,C + (1-x)\,D$ is the total payoff received by the entire society where $C(x)$ and $D(x)$ are the reward/penalty functions as defined by equations (1) and (2). This quantity is the so-called production function for the collective action of the society [26]. When the total payoff is zero, it is easy to think that nothing will happen and an equilibrium is reached. This is, however, not true.

For Pavlovian agents, analytical solutions of N-person games are possible [27]. When the cooperators receive the same total payoff as the defectors, then

$$x\,C(x) = (1-x)\,D(x). \tag{4}$$

If $C(x)$ and $D(x)$ are both linear functions of x, then this is a quadratic equation; if $C(x)$ and $D(x)$ are quadratic functions, then it is a cubic equation, and so on. The equation generally has up to two real solutions. If both solutions are in the interval $0 < x < 1$, then both equilibria are present. We will denote these equilibrium solutions x_1 and x_2, so that $0 < x_1 < x_2 < 1$. The initial cooperation probability (which is set as a constant and uniform across all the agents) is x_0.

The two solutions are different from each other in three important ways:

1. The solution at x_1 is a stable equilibrium (attractor) with respect to the aggregate cooperation proportion while the solution at x_2 is an unstable equilibrium (repulsor).

2. When $x_0 < x_2$, the solution converges toward x_1 as an oscillation while it stabilizes exactly in the $x_0 > x_2$ case.

3. Initial aggregate cooperation proportions of $x_0 > x_2$ do not result in the aggregate cooperation proportion converging to 1, as would be expected. This is because, for an individual agent that started off as a defector, there is always some likelihood that the agent will continue to defect. This probability is initially small but continues to increase if the agent is always rewarded for defecting. If the number of agents is sufficiently large and x_0 is not too close to 1, then there will be some agents that continue to defect until their cooperation probability reaches zero due to the successive rewards they have received, and these agents will defect forever. The exception is if you start off with the aggregate cooperation proportion equal to 1. Then no agent starts as a defector and there is no chance of any of them defecting in the steady state.

Substituting

$$C(x) = 1 - \frac{x}{L} \tag{5}$$

https://doi.org/10.25088/ComplexSystems.21.2.153

and

$$D(x) = -1 + \frac{x}{L} \tag{6}$$

into equation (4) from Figures 1 or 2, we obtain that the solutions x_{final} of our games will always converge to

$$x_{\text{final}} = L, \tag{7}$$

which is exactly true as demonstrated by the following simulations.

Start with the original problem: there are 100 agents and $L = 0.6$. The results of the simulation are shown in Figure 3. At whatever initial attendance we start, the attendance will wildly fluctuate around the value of L. Two extreme cases (total and zero initial attendances) are shown in Figure 3. The amplitude of fluctuations is about 20% of the total number of agents, like in the original presentation of the problem [1].

It is clear that these fluctuations are the consequence of the small number of agents. If we imagine a much larger bar, the fluctuations will be much smaller. This fact is demonstrated for the case of 10 000 agents as follows.

The result is always the same: minor fluctuations around the value of L. $L = 0.1$ in Figure 4, $L = 0.5$ in Figure 5, and $L = 0.9$ in Figure 6, for any initial attendance values.

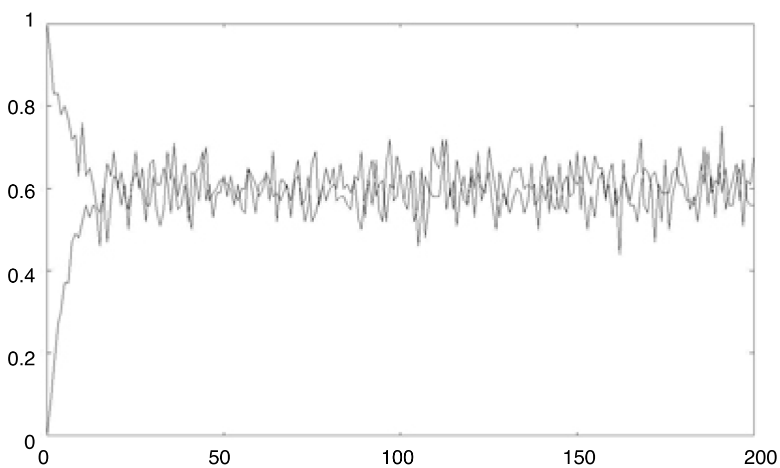

Figure 3. The El Farol Bar problem for 100 agents and $L = 0.6$. The initial ratio of bar-goers is 0 or 1.

https://doi.org/10.25088/ComplexSystems.21.2.153

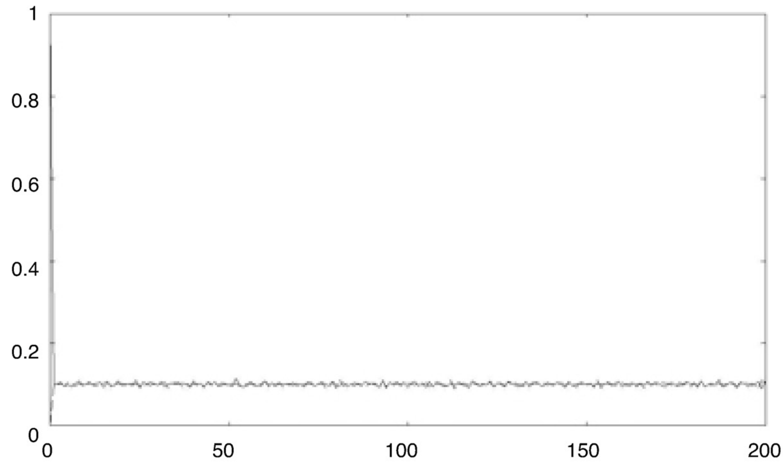

Figure 4. The El Farol Bar problem for 10 000 agents and $L = 0.1$. The initial ratio of bar-goers is 0 or 1.

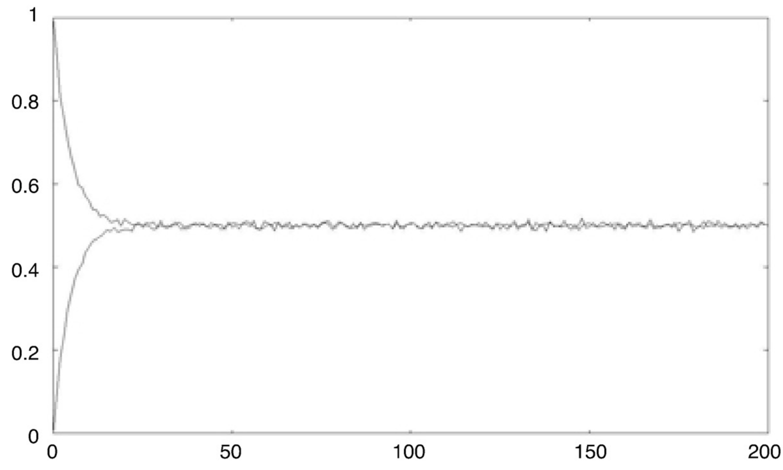

Figure 5. The El Farol Bar problem for 10 000 agents and $L = 0.5$. The initial ratio of bar-goers is 0 or 1.

https://doi.org/10.25088/ComplexSystems.21.2.153

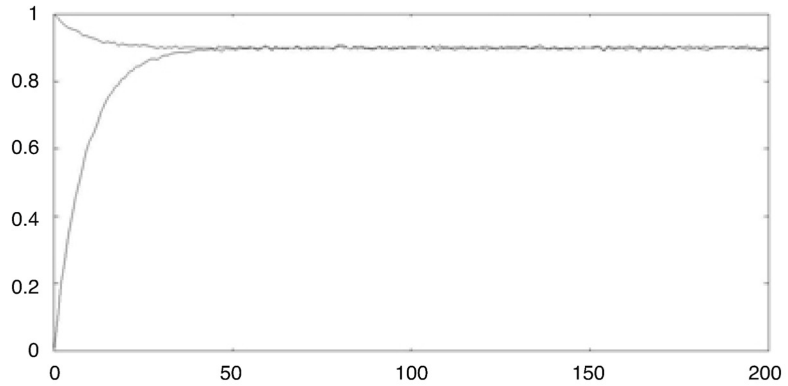

Figure 6. The El Farol Bar problem for 10 000 agents and $L = 0.9$. The initial ratio of bar-goers is 0 or 1.

5. Conclusion

The El Farol Bar problem can be solved as a special case of the N-person battle of the sexes or leader game for any initial attendance ratio and any value of the bar capacity. The result is always a fluctuating attendance around the value of L.

References

[1] W. B. Arthur, "Inductive Reasoning and Bounded Rationality," *The American Economic Review*, **84**(2), 1994 pp. 406–411. http://pluto.dma.unive.it/~paolop/ecomp/papers/arthur94.pdf.

[2] J. L. Casti, "Seeing the Light at El Farol," *Complexity*, **1**(5), 1995 pp. 7–10.

[3] R. Franke, "Reinforcement Learning in the El Farol Model," *Journal of Economic Behavior and Organization*, **51**(3), 2003 pp. 367–388. doi:10.1016/S0167-2681(02)00152-X.

[4] D. Challet, M. Marsili, and G. Ottino, "Shedding Light on El Farol," *Physica A*, **332**(1), 2004 pp. 469–482. doi:10.1016/j.physa.2003.06.003.

[5] D. Challet and Y.-C. Zhang, "Emergence of Cooperation and Organization in an Evolutionary Game," *Physica A*, **246**(3–4), 1997 p. 407. doi:10.1016/S0378-4371(97)00419-6.

[6] N. F. Johnson, P. M. Hui, D. Zheng, and C. W. Tai, "Minority Game with Arbitrary Cutoffs," *Physica A*, **269**(2–4), 1999 pp. 493–502. doi:10.1016/S0378-4371(99)00117-X.

[7] M. Hart, P. Jefferies, N. F. Johnson, and P. M. Hui, "Generalized Strategies in the Minority Game," *Physical Review E*, **63**(1), 2001 p. 017102. doi:10.1103/PhysRevE.63.017102.

[8] K. Lee, P. M. Hui, and N. F. Johnson, "The Minority Game with Different Payoff Functions: Crowd–Anticrowd Theory," *Physica A*, **321**(1–2), 2003 pp. 309–317. doi:10.1016/S0378-4371(02)01786-7.

[9] U. Wilensky. "NetLogo Minority Game Model." (2004) Center for Connected Learning and Computer-Based Modeling, Evanston, IL: Northwestern University.
http://ccl.northwestern.edu/netlogo/models/MinorityGame.

[10] N. F. Johnson, P. Jefferies, and P. M. Hui, *Financial Market Complexity*, Oxford, UK: Oxford University Press, 2003.

[11] A. C. C. Coolen, *The Mathematical Theory of Minority Games: Statistical Mechanics of Interacting Agents*, Oxford, UK: Oxford University Press, 2005.

[12] D. Challet, M. Marsili, and Y.-C. Zhang, *Minority Games*, Oxford, UK: Oxford University Press, 2005.

[13] J. Zhao, M. N. Szilagyi, and F. Szidarovszky, "An *N*-Person Battle of Sexes Game," *Physica A*, **387**(14), 2008 pp. 3669–3677.
doi:10.1016/j.physa.2007.09.053.

[14] M. N. Szilagyi and I. Somogyi, "Agent-Based Simulation of *N*-Person Games with Crossing Payoff Functions," *Complex Systems*, **17**(4), 2008 pp. 427–439. http://www.complex-systems.com/pdf/17-4-7.pdf.

[15] M. N. Szilagyi and Z. C. Szilagyi, "A Tool for Simulated Social Experiments," *Simulation*, **74**(1), 2000 pp. 4–10.
doi:10.1177/003754970007400101.

[16] S. Wolfram, *Cellular Automata and Complexity: Collected Papers*, Reading, MA: Addison-Wesley, 1994.

[17] R. Hegselmann and A. Flache, "Understanding Complex Social Dynamics: A Plea for Cellular Automata Based Modelling," *Journal of Artificial Societies and Social Simulation*, **1**(3), 1998 p. 3.
http://jasss.soc.surrey.ac.uk/1/3/1.html.

[18] K. S. Narendra and M. A. L. Thathachar, *Learning Automata: An Introduction*, Englewood Cliffs, NJ: Prentice Hall, 1989.

[19] A. Flache and M. W. Macy, "Weakness of Strong Ties: Collective Action Failure in a Highly Cohesive Group," *Journal of Mathematical Sociology*, **21**(1–2), 1996 pp. 3–28. doi:10.1080/0022250X.1996.9990172.

[20] S. S. Komorita, "A Model of the *n*-Person Dilemma-Type Game," *Journal of Experimental Social Psychology*, **12**(4), 1976 pp. 357–373.

[21] M. A. Nowak and R. M. May, "Evolutionary Games and Spatial Chaos," *Nature*, **359**, 1992 pp. 826–829. doi:10.1038/359826a0.

[22] E. L. Thorndike, *Animal Intelligence*, Darien, CT: Hafner, 1911.

[23] D. Kraines and V. Kraines, "Pavlov and the Prisoner's Dilemma," *Theory and Decision*, **26**(1), 1989 pp. 47–79. doi:10.1007/BF00134056.

[24] M. W. Macy, "Pavlov and the Evolution of Cooperation: An Experimental Test," *Social Psychology Quarterly*, **58**(2), 1995 pp. 74–87.

[25] A. Flache and R. Hegselmann, "Rationality vs. Learning in the Evolution of Solidarity Networks: A Theoretical Comparison," *Computational and Mathematical Organization Theory*, **5**(2), 1999 pp. 97–127. doi:10.1023/A:1009662602975.

[26] M. N. Szilagyi, “Quantitative Relationships between Collective Action and Prisoners’ Dilemma,” *Systems Research and Behavioral Science*, **17**(1), 2000 pp. 65–72. doi:10.1002/(SICI)1099-1743(200001/02)17:1<65::AID-SRES272>3.0.CO;2-U.

[27] M. N. Szilagyi, “An Investigation of *N*-Person Prisoners’ Dilemmas,” *Complex Systems*, **14**(2), 2003 pp. 155–174. http://www.complex-systems.com/pdf/14-2-3.pdf.