The N-Player Trust Game and its Replicator Dynamics

Hussein Abbass, Garrison Greenwood, and Eleni Petraki

Abstract—**Trust is a fundamental concept that underpins the coherence and resilience of social systems and shapes human behavior. Despite the importance of trust as a social and psychological concept, the concept has not gained much attention from evolutionary game theorists. In this paper, an N-player trust-based social dilemma game is introduced. While the theory shows that a society with no untrustworthy individuals would yield maximum wealth to both the society as a whole and the individuals in the long run, evolutionary dynamics show this ideal situation is reached only in a special case when the initial population contains no untrustworthy individuals. When the initial population consists of even the slightest number of untrustworthy individuals, the society converges to zero trusters, with many untrustworthy individuals. The promotion of trust is an uneasy task, despite the fact that a combination of trusters and trustworthy trustees is the most rational and optimal social state. This paper presents the game and results of replicator dynamics in a hope that researchers in evolutionary games see opportunities in filling this critical gap in the literature.**

Index Terms—**Trust, Evolutionary Game Theory, Trust Game, N-Person Trust Game**

I. Introduction

TRUST is the glue of a social system [24]. Despite its vital role in the society, the concept is absent from the evolutionary computation literature. When compared to the large number of papers published on the iterated prisoner dilemma [13], [6], [21], [2], [27], [23], there has not been any publication on trust. The contributions of this paper are three-fold. First, it aims to encourage more work on trust in the evolutionary computation and evolutionary game theory research areas. Second, it introduces a novel N-player trust game that assists researchers in understanding and analyzing the concept of trust using evolutionary game theory. The game is a social dilemma and generalizes the concept of trust, which is normally modelled as a sequential game, to a population of players that can play the game concurrently. Third, it presents the first theoretical analysis of the dynamics of the game using replicator dynamics.

A recent review on trust in social and psychological literature [30] demonstrated that there are abundant studies on the roles of trust [10], [11], [24] and its implications for social and human systems [12], [14], [17], [31], including ethical considerations [3] surrounding studies of trust and the means for influencing and shaping trust [5], [26]. Managers perceive how an understanding of how trust is formed can create opportunities to develop loyal customers [4] and improve relationships among employees and management [19], [33], [32]. Individuals perceive trust as a vulnerability [26] in which one person exposes himself to another by relying on the other person to make a decision on his/her behalf, as an opportunity [28] to create favorable outcomes, and as a source of unwanted uncertainty [16] that creates a relationship in which the trustee has power over the trusting party.

In the social and psychological literature, trust is described as playing two important roles. First, sociologically, it acts as a complexity-reduction mechanism [24], allowing individuals with limited cognitive capacities to manage the complex world they live in. Luhmann [24] showed that trust creates a positive feedback loop in a social system. As individuals use trust to manage complexity, relationships emerge, a process that increases complexity and thus creates more reinforcement and opportunities for trust to spread.

Second, psychologically, it acts as an ambiguity- and uncertainty-reduction mechanism for individuals [9], [11]. Deutsch [11] argues that a trusting situation occurs when a truster perceives that a situation has one of two potential outcomes, where one is perceived to have a negative valency of greater absolute value than the positive valency attributed to the second. However, which outcome will occur depends on the trustee. If the truster chooses to proceed, the truster is said to trust the trustee; otherwise, the truster distrusts the trustee. Context influences the perceived levels of both trustworthiness and risk; as the risk of trusting changes from one context to another, the decision to trust changes as well. Therefore, trust and risk are tightly coupled concepts.

Trust games are sequential in nature. The truster must first decide whether to trust the trustee. If the decision is to trust the trustee, the trustee must then decide whether to be trustworthy or not. Each decision can be a binary decision or be considered as part of a continuum representing the degree of either trust or trustworthiness. Although the proposed game relies on binary decision-making, it can easily be generalized to include non-binary decisions. Neuroeconomic experiments use two main forms of trust games [1], [9], [16]. We will call them TG1 and TG2 for trust game 1 and 2, respectively.

TG1 [22] is a non-zero-sum game, and the truster must choose whether to trust the trustee. If the truster chooses not to trust the trustee, both players receive \$5. If the truster chooses to trust the trustee, then the trustee must decide whether to be trustworthy or not. If the trustee chooses to be trustworthy, the truster and trustee receive \$10 and \$15, respectively. If the trustee chooses to defect, the truster gets nothing, whereas the trustee gets \$25.

Hussein Abbass is with the School of Engineering & IT, University of New South Wales, Canberra-Australia, (e-mail: h.abbass@adfa.edu.au).

Garrison Greenwood is with the Electrical and Computer Engineering Department, Portland State University, Oregon - USA (e-mail: greenwd@pdx.edu).

Eleni Petraki is with the Faculty of Arts and Design, University of Canberra Canberra, Australia (e-mail: Eleni.Petraki@canberra.edu.au).

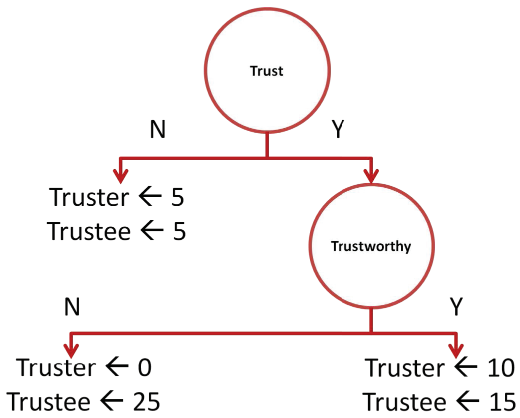

Fig. 1. Game tree for a 2-player TG1.

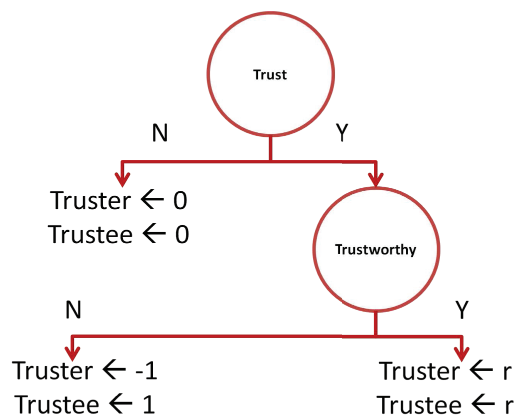

Fig. 2. Game tree for a general 2-player TG1G.

In this game, a rational truster would be indifferent to the two choices under the assumption that the trustee's probability of being trustworthy is 0.5. Therefore, the expected return for the truster, regardless of the decision to trust or not, is $5. The trustee, however, benefits more if the truster chooses to trust the trustee. In this case, in a one-off situation, a rational trustee would choose to defect to maximize his or her own return. However, over repeated iterations, the trustee has an incentive to be trustworthy because being so will yield maximum returns to both the trustee and truster. The game tree for this game can be seen in Figure 1.

TG1 can be generalized, we call it TG1G, to the game tree shown in Figure 2. An early version of this general form was presented in [7] and the general version was then analyzed in [25].

TG2 [1] is also a non-zero-sum game. Both the truster and trustee start with an endowment of $12. The truster makes the first move by deciding how much money to transfer to the trustee. The money is tripled on the way to the trustee. The trustee then must decide how much money to transfer back to the truster. If X is the amount of money the truster sent to the trustee and Y is the amount of money the trustee sends back to the truster, then the truster will end up with a balance of 12 - X + Y while the trustee ends up with a balance of 12 + 3X - Y. Both the truster and trustee end up with the same balance

TABLE I
UTILITY MATRIX FOR A 2-PLAYER TRUST GAME.

	Truster	Trustee
Starting Balance	12	12
Pay	X	Y
Receive	Y	$3X$
Net Wealth	$12 - X + Y$	$12 + 3X - Y$

when Y = 2X. In this game, X is a measure of trust and Y is a measure of trustworthiness. This game is represented in the following table.

Neither TG1 nor TG2 create a complete social dilemma. If the truster chooses to trust the trustee, then the total wealth, independent of whether the trustee is trustworthy or not, is $25 in TG1 and $24 + 2X$ in TG2. The social dilemma, therefore, is reliant only on the truster's decision. However, in TG1G, the game is a social dilemma when $0 < r < 1$.

II. N-PLAYER TRUST GAME

Assume N players. Each player must make two decisions in advance. The first is to decide whether to be trustworthy or not. The second is to decide whether to govern (be a governor or a trustee) or be governed (be a citizen or a truster). We will denote the former as G and the latter as C. Assume the number of players that decided to be type C is x_1 and the number of players that decided to be type G is $x_2 + x_3$; such that, $N = x_1 + x_2 + x_3$.

In real-world settings, the role of the truster and trustee are defined *a priori*. For example, an investor chooses to invest, while a financial planning advisor chooses to become one before the two actors decide to enter into a trusting relationship. This type of decision we refer to as a social choice. The second type of decision normally decided *a priori* as well is whether the trustee will be trustworthy or not. The reason it is 'normally' decided *a priori* is that it largely depends on the behavioral attitude and value system of the trustee. We acknowledge that context influences this behavioral attribute but we conjecture that this decision is fundamentally a core behavioral attribute of the agent. We call this decision a behavioral choice.

In summary, it is reasonable to assume that the two types of decisions related to social and behavioral choices can be done *a priori*.

A player of type C pays tv to the government, where tv denotes the trusted value. The dynamics of the game is in general independent of the value of tv, which can be set to 1. However, we maintain tv to allow flexibility in adopting the game to different contexts. With x_1 players of type C, the total money that is sent to the government is $x_1 \cdot tv$. Each player of type G receives $(x_1 \cdot tv)/(x_2 + x_3)$. Assume x_2 players of type G decide to be trustworthy, while x_3 decide to not be trustworthy. A player in the x_2 population returns to the citizens a multiplier of $R1$ of what was received and keeps the same amount for himself, with $R1 > 1$. A player in the x_3 population returns nothing to the citizens and keeps for himself a multiplier of $R2$ of what was received, where $R1 < R2 <$

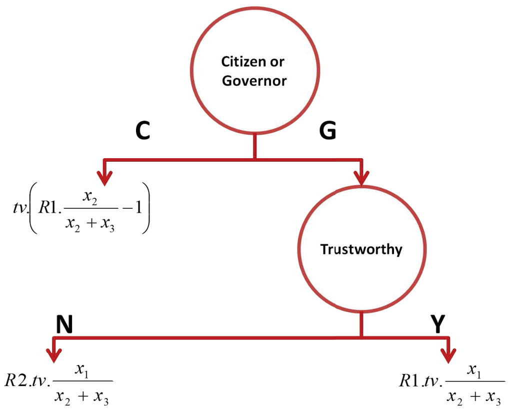

Fig. 3. Game tree for the proposed N-player trust game.

$2R1$. The payoff matrix for this game can then be represented as shown in Table II with the following constraints:

$$1 < R1 < R2 < 2R1 \tag{1}$$

$$N = x_1 + x_2 + x_3 \tag{2}$$

The game tree is presented in Figure 3 for an arbitrary number of players.

Theorem 1: $x_3 = N$ is a Nash Equilibrium.

Proof. The order of the net individual wealth in the table from left to right is a monotonically increasing sequence. A player deciding to play according to any strategy on the right-hand side will yield a better net individual wealth than any of the strategies on the left-hand side. The best net individual wealth occurs for players in x_3. Any rational player will choose to be a governor and untrustworthy. Thus, the Nash equilibrium occurs when all players choose to be governors and untrustworthy. ■

The total combined wealth CW of the population is

$$CW = \begin{cases} x_1 \cdot tv\left(\frac{2 \cdot R1 \cdot x_2}{x_2+x_3} + \frac{R2 \cdot x_3}{x_2+x_3} - 1\right) & \text{if } x_2 + x_3 \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Theorem 2: The maximum of CW occurs when $x_2 = 1$ and $x_3 = 0$.

Proof. Given the constraints on $R1$ and $R2$, CW has a lower bound of zero regardless of the value of $x_2 + x_3$. Therefore, to maximize CW, we only need to focus on the case when $x_2 + x_3 \geq 1$. Substituting $x_3 = (x_2 + x_3) - x_2$ into Eq. (3),

$$\begin{aligned} CW &= x_1 \cdot tv \cdot \left(\frac{2R1 \cdot x_2}{x_2 + x_3} - 1 + R2 - \frac{R2x_2}{x_2 + x_3}\right) \\ &= x_1 \cdot tv \cdot \left(\frac{(2R1 - R2) \cdot x_2}{x_2 + x_3} + (R2 - 1)\right) \\ &= (N - (x_2 + x_3)) \cdot tv \cdot \\ &\qquad \left(\frac{(2R1 - R2) \cdot x_2}{x_2 + x_3} + (R2 - 1)\right) \end{aligned}$$

Recall that $R2 > (2R1 - R2) > 0$, N is a constant, and $x_2 + x_3 \geq 1$, the maximum of the function will occur when $x_2 + x_3$ is at a minimum ($x_2 + x_3$ appears twice, once with a negative sign in the numerator and once in the denominator) while $x_3 = 0$ is the maximum for $\frac{x_2}{x_2+x_3}$; therefore, $x_2 + x_3 = 1 \quad \Rightarrow \quad x_2 = 1$. ■

Theorem 3: $x_1 = N - 1$ and $x_2 = 1$ is Pareto Optimal.

Proof. The single x_2 player can switch to an x_3 player or an x_1 player; either change reduces the x_1 player payoff to zero. All investments are split among the x_2 and x_3 players so any x_1 player that changes to an x_2 or x_3 player reduces the payoff to the single x_2 player. The proof follows that any player unilaterally switching roles reduces the payoff to an x_1 player. ■

Our investigations indicate other Pareto Optimal solutions exist, but the one in Theorem 3 is known as a social welfare maximizer or socially optimal solution because it maximizes the total utility for the population.

If every player chooses to be a governor and untrustworthy, the net individual wealth is zero. In contrast, if a player chooses to be in the x_1 population, he or she can lose all of their money if all of the governors are untrustworthy. Would a single trustworthy governor emerge in this population to maximize the combined wealth of the population?

The evolutionary behavior of a population playing the trust game can be studied using replicator dynamics [20]. Let y_k be the frequency of the x_k players in an infinitely large population with $\sum_k y_k = 1$. Then the time evolution of y_k is given by the differential equation

$$\dot{y}_k = y_k \cdot \left(f_k - \bar{f}\right)$$

where f_k is the expected fitness of an individual playing strategy k at time t and $\bar{f}$ is the mean population fitness. Here, fitness and net wealth are equivalent. The number of copies of a strategy increases if $f_k > \bar{f}$ and decreases if $f_k < \bar{f}$. We can calculate $\bar{f}$ as follows

$$\bar{f} = \frac{y_1 \cdot y_2 \cdot tv\,(2 \cdot R1 - 1) + y_1 \cdot y_3 \cdot tv \cdot (R2 - 1)}{(y_2 + y_3)}$$

The three replicator equations (Note that $y_2 + y_3 = 1 - y_1$) are

$$\begin{aligned} \dot{y}_1 = \frac{y_1^2 \cdot tv}{1 - y_1} \left(y_2 \left(1 - 2 \cdot R1\right) + y_3 \cdot \left(1 - R2\right)\right) + \\ \frac{y_1 \cdot tv}{1 - y_1} \left(y_2 \left(R1 - 1\right) - y_3\right) \end{aligned}$$

$$\dot{y}_2 = \frac{y_1 \cdot y_2 \cdot tv}{1 - y_1} \cdot \left(y_2 \left(1 - 2 \cdot R1\right) + y_3 \left(1 - R2\right) + R1\right)$$

$$\dot{y}_3 = \frac{y_1 \cdot y_3 \cdot tv}{1 - y_1} \cdot \left(y_2 \left(1 - 2 \cdot R1\right) + y_3 \left(1 - R2\right) + R2\right)$$

TABLE II
UTILITY MATRIX FOR A N-PLAYER TRUST GAME.

	Player in the x_1 population	**Player in the x_2 population**	**Player in the x_3 population**
Pay	tv	$R1 \cdot tv \cdot \frac{x_1}{x_2+x_3}$	0
Receive	$R1 \cdot tv \cdot \frac{x_2}{x_2+x_3}$	$2 \cdot R1 \cdot tv \cdot \frac{x_1}{x_2+x_3}$	$R2 \cdot tv \cdot \frac{x_1}{x_2+x_3}$
Net Wealth	$tv \cdot (R1 \cdot \frac{x_2}{x_2+x_3} - 1)$	$R1 \cdot tv \cdot \frac{x_1}{x_2+x_3}$	$R2 \cdot tv \cdot \frac{x_1}{x_2+x_3}$

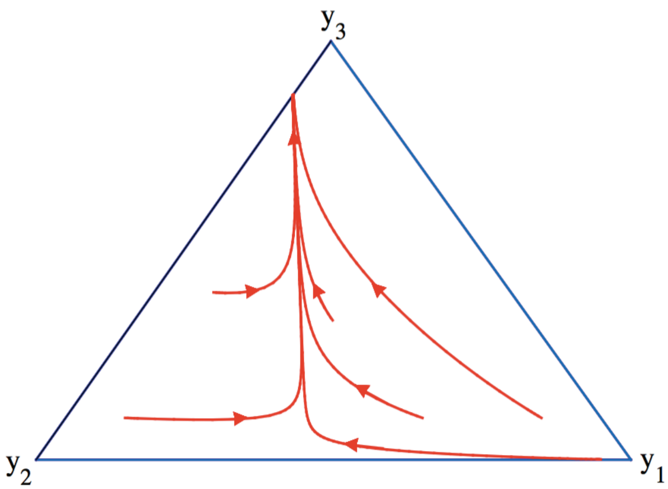

Fig. 4. A 2-simplex showing the time evolution for a game with $R1 = 6$, $R2 = 8$, $tv = 10$, and different initial distributions of y_1, y_2 and y_3.

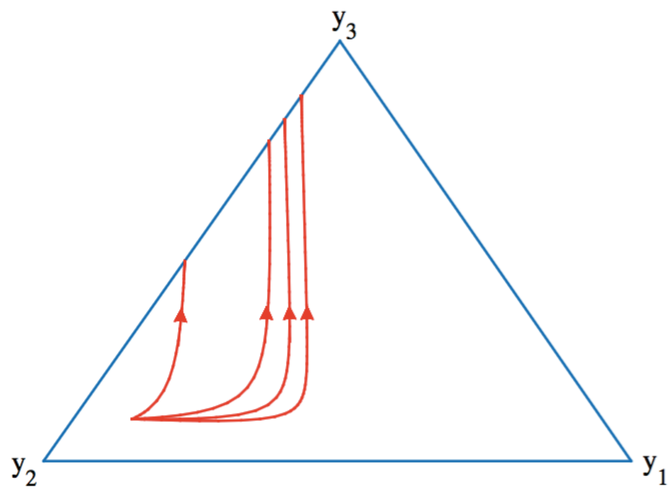

Fig. 5. A 2-simplex showing the time evolution for a game with $tv = 10$ and different $R1$ and $R2$ values ($R1 < R2$). Values increase from left to right with $R1 = 1.5, R2 = 2.9$ for the far left trajectory to $R1 = 6, R2 = 8$ for the far right trajectory. Initial distribution is $y_1(0) = 0.1$, $y_2(0) = 0.8$, and $y_3(0) = 0.1$.

Figure 4 shows the population evolution for various initial player distributions. The replicator equations predict there is a critical ratio of x_1 to x_2 players that acts as an attractor where players rapidly switch strategy to become untrustworthy. However, untrustworthy players do not completely take over the population because a certain proportion always remains trustworthy even when there are a few or no investors left. Without investors—i.e., $y_1 = 0$—each replicator equation has the form $\dot{y}_k = 0$, which is a fixed point. Importantly, when there are no longer any investors, the net worth of all remaining players is zero. Consequently, there is no incentive to switch strategies and a steady-state condition exists.

The bottom trajectory in Figure 4 shows an extreme case that is highly favorable to a small number of trustworthy players: a population composed almost entirely of investors with almost no untrustworthy players. Under these circumstances, the net worth of a trustworthy player is much higher than that of an investor; therefore evolution should favor the trustworthy players. Untrustworthy players have an even higher net worth, but they are initially rare in the population. The replicator equations initially predict very rapid growth in trustworthy players with a corresponding plummet in investors. However, the inevitable increase of untrustworthy players reverses the trustworthy player growth. Even in this extreme case, the steady-state condition is reached.

The replicator equations predict interesting behavior in the trust game. Regions to the left of the attractor have low C to G ratios. Since investments are split among G players higher returns go to x_1 players if most G players are trustworthy. The replicator equations predict x_2 players will mostly switch to x_1 players. Conversely to the right of the attractor there is a high C to G ratio. Under those circumstances there is a strong temptation to be untrustworthy and the replicator equations predict a sharper rise in x_3 players. Near the attractor the returns to x_1 and x_2 players are roughly the same so there is little incentive to switch from x_2 to x_1 or vica versa. The maximum return is obtained by becoming untrustworthy and the replicator equations predict virtually all strategy changes are to x_3 players. Eventually the trajectories intersect the $y_2 - y_3$ line where there are no investors. At that point all replicator equations are of the form $\dot{y}_k = 0$ and a fixed point is reached. This makes perfect sense because, without any investors, there is no incentive to switch strategies.

Figure 5 shows the effect of different $R1$ and $R2$ values (but with $R1 < R2$). A fixed point is still always reached, but the lower the values, the more trustworthy players in the final population. The reason is lower $R1$ and $R2$ values cause the y_1 players to go extinct quicker, which limits the growth of untrustworthy players. Notice the similarity of these trajectories to those in Figure 4, which suggests the presence of an attractor. However, although all attractors are oriented in the same direction, their location in the 2-simplex depends on the $R1$ and $R2$ values.

Theorem 4: $y_3 = 0$, $y_1 = \frac{R1-1}{2R1-1}$ and $y_2 = \frac{R1}{2R1-1}$ is a

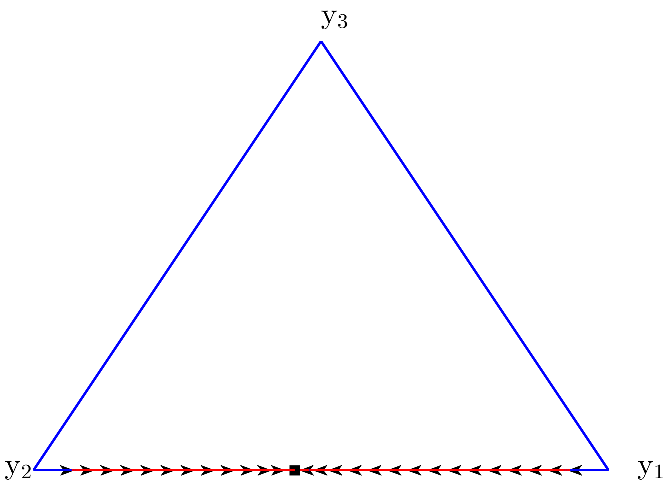

Fig. 6. A 2-simplex showing the time evolution for a game with $tv = 10$, $R1 = 6$, $R2 = 8$ with different ratios of y_1 and y_2 while maintaining $y_3 = 0$ and $y_1 + y_2 + y_3 = 1$.

fixed point.

Proof. The above fixed point is a non-trivial solution for the replicator equations as a homogeneous system.■

Figure 6 shows the fixed point discussed in Theorem 4. Any starting point on the $y_1 - y_2$ axis, where $y_1 + y_2 + y_3 = 1$ and $y_3 = 0$ would converge to a fixed point on this axis. The fixed point in this example, where $R1 = 6$ is $y_1 = 0.454$ and $y_2 = 0.545$. This fixed point is shown with a black square in Figure 6.

In summary, the replicator equations predict a rapid growth of untrustworthiness in the population, leading to the eventual extinction of investors. However, a fraction of the population always remains trustworthy, even in the absence of investors. This predicted steady-state outcome is independent of the initial player distribution, but the ratio of trustworthy to untrustworthy players in the final population is dependent on the $R1$ and $R2$ values.

The above study and the literature review demonstrate that the area of modelling trust in evolutionary game theory is in its infancy. Many opportunities exist to study this fascinating topic. For example, similar to Yeh and Yang [34], one can study the dynamics of the game on a social network, or use genetic algorithms to evolve strategies for the game [8]. Many of the assumptions that have been relaxed for the iterated prisoner dilemma still hold in the trust games discussed above. For example, one can study the impact of memory size on the game [23] and the impact of memory on the cognitive resources of the agents [15]. Last, but not least, opportunities exist on the study of the fitness landscape of the trust game [18], [29], which can reveal insight into the level of hardness or difficulties in discovering novel strategies, especially in the context of mixed strategy.

III. Conclusion

Trust is a fundamental concept for social systems and individuals alike. Previous games of trust are limited to 2 players. Previous work in neuroeconomics introduced basic forms of two-player games to model trust and described the relationships between trusting decisions and human neural functions. These games do not create a social dilemma where the total social wealth depends on the decisions of both truster and trustee, and they do not generalize to multiple players. The field of evolutionary computation has seen little or no work on trust, despite the socioeconomic and psychological significances of the concept.

In this work, we introduce a new N-player trust game that can model trust decisions among many players. It creates a social dilemma in which individuals who attempt to maximize their own benefits in the short run maximize neither the society's social wealth nor their own benefits in the long run. The results revealed that, while the optimal solution for the population exists when N-1 players choose to be trusters and the remaining players choose to be trustees, this optimal solution is a needle in a haystack, causing the evolutionary dynamics to consistently converge to a population without any trusters and with a combination of trustworthy and untrustworthy individuals. The exception to this phenomenon occurs when the initial population is free of untrustworthy players. The proposed game shows that trust is an uneasy but worthwhile concept.

Acknowledgement

Portions of this work was funded by the Australian Research Council Discovery Grant number DP140102590.

References

[1] R. Adolphs, "Trust in the brain," *Nature neuroscience*, vol. 5, no. 3, pp. 192–193, 2002.
[2] D. Ashlock, E. Y. Kim, and N. Leahy, "Understanding representational sensitivity in the iterated prisoner's dilemma with fingerprints," *IEEE Transactions on Systems, Man and Cybernetics - Part C*, vol. 36, no. 4, 2006.
[3] A. Baier, "Trust and antitrust," *Ethics*, pp. 231–260, 1986.
[4] D. Ball, P. Simões Coelho, and A. Machás, "The role of communication and trust in explaining customer loyalty: An extension to the ecsi model," *European Journal of Marketing*, vol. 38, no. 9/10, pp. 1272–1293, 2004.
[5] A. W. Brooks, H. Dai, and M. E. Schweitzer, "Im sorry about the rain! superfluous apologies demonstrate empathic concern and increase trust," *Social Psychological and Personality Science*, vol. 5, no. 4, pp. 467–474, 2014.
[6] P. J. Darwen and X. Yao, "Speciation as automatic categorical modularization," *Evolutionary Computation, IEEE Transactions on*, vol. 1, no. 2, pp. 101–108, 1997.
[7] P. Dasgupta, "Trust as a commodity," *Trust: Making and breaking cooperative relations*, vol. 4, pp. 49–72, 2000.
[8] O. E. David, H. J. van den Herik, M. Koppel, and N. S. Netanyahu, "Genetic algorithms for evolving computer chess programs," *Evolutionary Computation, IEEE Transactions on*, vol. 18, no. 5, pp. 779–789, 2014.
[9] M. R. Delgado, R. H. Frank, and E. A. Phelps, "Perceptions of moral character modulate the neural systems of reward during the trust game," *Nature neuroscience*, vol. 8, no. 11, pp. 1611–1618, 2005.
[10] M. Deutsch, "Cooperation and trust: Some theoretical notes." in *Proceedings of Nebraska Symposium on Motivation*, M. R. Jones, Ed. Univer. Nebraska Press, 1962.
[11] ——, *The resolution of conflict: Constructive and destructive processes*. Yale University Press, 1977.
[12] E. Fehr and U. Fischbacher, "The nature of human altruism," *Nature*, vol. 425, no. 6960, pp. 785–791, 2003.
[13] D. B. Fogel, "Evolving behaviors in the iterated prisoner's dilemma," *Evolutionary Computation*, vol. 1, no. 1, pp. 77–97, 1993.
[14] D. Gambetta *et al.*, "Can we trust trust," *Trust: Making and breaking cooperative relations*, vol. 2000, pp. 213–237, 2000.

[15] A. Ghoneim, G. W. Greenwood, and H. Abbass, "Distributing cognitive resources in one-against-many strategy games," in *Evolutionary Computation (CEC), 2013 IEEE Congress on*. IEEE, 2013, pp. 1387–1394.

[16] K. Giffin, "The contribution of studies of source credibility to a theory of interpersonal trust in the communication process." *Psychological bulletin*, vol. 68, no. 2, p. 104, 1967.

[17] F. S. Grodzinsky, K. W. Miller, and M. J. Wolf, "Developing automated deceptions and the impact on trust," *Philosophy & Technology*, pp. 1–15, 2014.

[18] J. He, T. Chen, and X. Yao, "On the easiest and hardest fitness functions," *Evolutionary Computation, IEEE Transactions on*, vol. 19, no. 2, pp. 295–305, 2015.

[19] J. F. Helliwell and H. Huang, "Well-being and trust in the workplace," *Journal of Happiness Studies*, vol. 12, no. 5, pp. 747–767, 2011.

[20] J. Hofbauer and K. Sigmund, "Evolutionary game dynamics," *Bulletin of the American Mathematical Society*, vol. 40, no. 4, pp. 479–519, 2003.

[21] H. Ishibuchi and N. Namikawa, "Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing," *Evolutionary Computation, IEEE Transactions on*, vol. 9, no. 6, pp. 552–561, 2005.

[22] F. Krueger, K. McCabe, J. Moll, N. Kriegeskorte, R. Zahn, M. Strenziok, A. Heinecke, and J. Grafman, "Neural correlates of trust," *Proceedings of the National Academy of Sciences*, vol. 104, no. 50, pp. 20 084–20 089, 2007.

[23] J. Li and G. Kendall, "The effect of memory size on the evolutionary stability of strategies in iterated prisoner's dilemma," *Evolutionary Computation, IEEE Transactions on*, vol. 18, no. 6, pp. 819–826, 2014.

[24] N. Luhmann, H. Davis, J. Raffan, and K. Rooney, *Trust; and, Power: two works by Niklas Luhmann*. Wiley Chichester, 1979.

[25] N. Masuda and M. Nakamura, "Coevolution of trustful buyers and cooperative sellers in the trust game," *PloS one*, vol. 7, no. 9, p. e44169, 2012.

[26] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Academy of management review*, vol. 20, no. 3, pp. 709–734, 1995.

[27] S. Mittal and K. Deb, "Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives," *Evolutionary Computation, IEEE Transactions on*, vol. 13, no. 3, pp. 554–565, 2009.

[28] E. W. Morrison and S. L. Robinson, "When employees feel betrayed: A model of how psychological contract violation develops," *Academy of management Review*, vol. 22, no. 1, pp. 226–256, 1997.

[29] M. Munoz, M. Kirley, S. K. Halgamuge *et al.*, "Exploratory landscape analysis of continuous space optimization problems using information content," *Evolutionary Computation, IEEE Transactions on*, vol. 19, no. 1, pp. 74–87, 2015.

[30] E. Petraki and H. Abbass, "On trust and influence: A computational red teaming game theoretic perspective," in *Computational Intelligence for Security and Defense Applications (CISDA), 2014 Seventh IEEE Symposium on*. IEEE, 2014, pp. 1–7.

[31] D. S. Reina and M. L. Reina, "Trust and betrayal in the workplace building effective relationships in your organization," *Advances in Developing Human Resources*, vol. 2, no. 1, pp. 121–121, 2000.

[32] M. D. Spector and G. E. Jones, "Trust in the workplace: Factors affecting trust formation between team members," *The Journal of social psychology*, vol. 144, no. 3, pp. 311–321, 2004.

[33] C. Wells, "Trust, gender, and race in the workplace," *Journal of Social Behavior and Personality*, vol. 16, no. 1, pp. 115–126, 2001.

[34] C. Yeh and C. Yang, "Social networks and asset price dynamics," *Evolutionary Computation, IEEE Transactions on*, vol. 19, no. 3, pp. 387–399, 2015.