Revisiting the Evolution of Cooperation in Complex Societies

Extended Abstract

Jonathan Skaggs
Brigham Young University
Provo, UT, USA
jbskaggs12@gmail.com

Jacob W. Crandall
Brigham Young University
Provo, UT, USA
crandall@cs.byu.edu

ABSTRACT

Cooperation is a foundation of human societies. Nevertheless, aggression and exploitation persist even in highly interconnected and cooperative societies. This coexistence of cooperation and aggression is not captured by many models. In this work, we demonstrate that cooperation and aggression can coexist when models consider asymmetric power, interdependence, and evolving social structure. Specifically, we examine the evolution of cooperation and the emergence of aggression within the Junior High Game (JHG), a model of society that embeds agents within a complex system that embodies these attributes. We show that, within this model, the information scope experienced by members of society shapes the emergence, stability, and form of cooperation and aggression present in society. Expanding the information scope stabilizes cooperation by enabling reciprocity and coalition-like protection, but also enables new forms of aggression that exploit network structure. In particular, within this model, richer information can legitimize aggressive norms in which individuals and groups systematically target weaker entities. These results suggest that increased social awareness does not eliminate aggression, but transforms how it emerges and persists.

KEYWORDS

Complex Systems, Emergent Behavior, Game Theory

ACM Reference Format:
Jonathan Skaggs and Jacob W. Crandall. 2026. Revisiting the Evolution of Cooperation in Complex Societies: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/OUAB7380

1 INTRODUCTION

Cooperation is a defining feature of human societies [4, 30], enabling compromise, economic growth, collective action, and social cohesion. Given this importance, extensive research has been devoted to understanding why and when people cooperate. Notably, classical models–including public goods games, prisoner's dilemmas, and extensions incorporating repetition, spatial structure, and networked interaction (e.g., [1, 3, 5, 8, 18–23, 26–28])–have been used to provide foundational insights into cooperation. Studies using these models often point to conditions under which society as a whole converges to cooperative or non-cooperative behavior.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/OUAB7380

And yet, aggression persists [6, 32] along-side cooperation. Exploitation, coercion, and opportunistic predation appear even in highly interconnected, norm-governed societies. This coexistence is not an anomaly, but a recurring pattern across social, economic, and organizational systems. On the international stage, trade (cooperation) among countries flourishes, but aggression, including invasion of other countries, continues. Businesses engage in mutually advantageous collaborations, but sometimes weaponize litigation to intimidate, delay, or otherwise stymie competitors. Finally, while individuals and groups often help each other, violent and aggressive behavior still permeates societies to varying degrees. These realities demonstrate that cooperation and aggression tend to coexist–simple convergence to cooperative or non-cooperative behavior as identified in many models is not the norm.

To more effectively understand the emergence of aggression alongside the evolution of cooperation, models incorporating properties of complex societies [13, 15, 16] should be used. In complex societies, global outcomes arise from local strategic choices, and cannot be predicted from individual behavior alone [2, 16]. Local actions cascade through networks, amplifying small incentives into large-scale norms [7, 10, 17, 31]. These complexities are not well-modeled in classical social dilemmas, which seemingly limits the applicability of insights they provide in regards to the evolution of cooperation. Thus, in this work, we re-examine the evolution of cooperation within a more complex societal model.

2 A MODEL OF COMPLEX SOCIETIES

We examine the evolution of cooperation and aggression in the Junior High Game (JHG) [25], a strategic network game encoding key properties of complex societies [9, 13–16]: asymmetric power, finite resources, interdependence, and evolving interaction networks. Rather than binary cooperate-or-defect choices, agents in the JHG allocate limited resources to individuals in society, choosing the degree to which they cooperate with, attack, or ignore each member. They also consider how much resources to expend on defending themselves against possible attacks from others. Power asymmetry and evolving interaction networks naturally arises from historical play of the JHG, often leading to (potentially competing) coalitions that persist, evolve, and dissolve over time.

Following Axelrod's study of the evolution of cooperation in the iterated prisoner's dilemma [1], we analyze strategies in the JHG with respect to evolutionary stability via replicator dynamics [12, 29]. That is, we enumerate strategies in the JHG, and then analyze the ability of each strategy to both invade other strategies and to avoid invasion. Specifically, we observe via replicator dynamics the ability of a strategy, employed by a small number of agents, to invade a society whose members employ an alternative strategy.

Table 1: Summary of primary results.

Information Scope	Key Result
Societal Response	The always-defend strategy (keep all tokens) dominates
Direct Interaction	Cooperative reciprocators thrive, but aggressive parasites persist
Indirect Interaction	Cooperation is the norm, but limited strategic exploitation of weaker individuals is often advantageous (depending on societal norms)

Table 2: Summary of considered strategies for allocating extra tokens in the direction-interaction scope.

Strategy	Description
Random	Randomly allocates extra tokens both positively and negatively
Cooperative	Randomly allocates extra tokens positively
Aggressive	Randomly allocates extra tokens negatively
Preferential	Positively allocates extra tokens to random individuals with a preference toward more influential players
Strengthening	Positively allocates extra tokens to random individuals with a preference toward friends
Defensive	Keeps all extra tokens tokens
Parasite	Exploits others by randomly selecting a target and attacking it with all its tokens until it is sufficiently weak

In these evaluations of evolutionary stability in the JHG, we consider the role of *information scope*. Information scope defines how much individuals can observe the behavior of other agents in society as well as their ability to control who they interact with. We examine agent strategies in three different information scopes: (1) societal-response (agents react to aggregate outcomes, approximating public goods games [20]), (2) direct-interaction (agents track bilateral histories, enabling reciprocity), and (3) indirect-interaction (agents observe triadic structure, connecting to structural balance theory [11] and indirect reciprocity [18]). Regardless of information scope, agents still interact using the same underlying JHG dynamics, but these different scopes alter agents' abilities to directly interact with others and observe their behaviors.

3 OVERVIEW OF RESULTS

A high-level summary of our results in each information scope is summarized in Table 1. In the societal-response scope, an agent can only observe and react to society's aggregate behavior. Thus, this scope has similarity to public good games. Unsurprisingly, cooperation collapses under this scope. Mirroring free-riding in public good games, the always-defend strategy, in which a player keeps all of its tokens, is evolutionarily stable and invades all other strategies we considered in this scope.

In the direct-interaction scope, agents track partner-specific histories. This enables tit-for-tat-like reciprocity that has been shown to be effective in iterated prisoners dilemmas [1]. In this scope, agents can reward cooperation from others by reciprocating tokens, defend against attacks, and retaliate against aggressors on a player-by-player basis. However, agents must allocate their finite resources across all players, meaning strategies must define how to allocate tokens given given excess or shortfalls. In addition to variations on reciprocal strategies (Table 2), we also consider a *Parasite* strategy which, unlike the reciprocal strategies, dedicates all tokens to opportunistic exploitation, attacking weak targets until extraction becomes unprofitable, then shifting victims.

Figure 1a shows evolutionary stability among pairwise interactions. Three strategies emerge as evolutionarily stable: Cooperative, Preferential, and Parasite. Cooperative and Preferential reciprocators achieve the highest societal productivity, validating Axelrod's insight that reciprocity stabilizes cooperation. However, Parasites are not eliminated. This strategy stabilizes at 30–40% of the population, creating a stable polymorphism where cooperation and exploitation coexist indefinitely. Figure 1b-c shows that mixed societies without parasites evolve toward cooperative equilibria, while those including parasites sustain diverse strategies in stable coexistence.

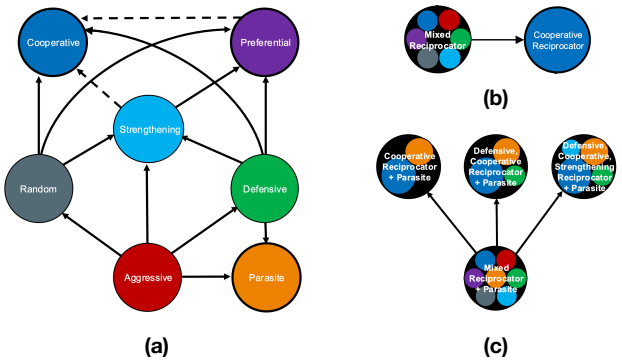

Figure 1: Stability flow diagrams given direct-interaction scope. (a) Evolutionary dynamics given pairwise comparisons. Directed edges represent successful invasions: an arrow from node A to node B indicates that a small group of invaders using strategy B can out-compete residents using strategy A. (b-c) Evolutionary dynamics when beginning from a mixed society with and without parasites.

In the indirect-interaction scope, agents can observe interactions between other players, thus allowing for indirect reciprocity and other strategies related to structural balance. Here, societies achieve high productivity through cooperation and norm consolidation, yet outcomes depend critically on initial conditions and exhibit pressure toward norm adoption. Interestingly, under certain initial conditions (norms), limited strategic exploitation of weaker individuals is often advantageous. Hence, societies tend to converge to somewhat cooperative strategies with occasional exploitation, a phenomena that reflects human behavior in the JHG [24]. Notably, exploiter-dominated societies often achieve the highest productivity, revealing a tension between collective welfare and the persistence of norm-enabled aggression.

4 CONCLUSION

Using the JHG, we demonstrate that Axelrod's insights regarding the evolution of cooperation generalize to more complex societies, but with critical qualifications. At societal-response scope, cooperation is fragile (free-riding dominates). At direct-interaction scope, reciprocity stabilizes cooperation, but coexists with stable aggressive (highly non-cooperative) strategies. At indirect-interaction scope, societies exhibit norm consolidation, with exploiter strategies often prevailing despite, or because of, their coordination through majority pressure. These results suggest that increased social awareness does not eliminate aggression, but transforms how it emerges and persists in otherwise cooperative societies.

REFERENCES

[1] Robert Axelrod. 1984. *The Evolution of Cooperation.* Basic Books, New York, NY, USA.

[2] Robert Axelrod. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration.* Princeton University Press, Princeton, NJ, USA.

[3] Alex J. Bladon, Tobias Galla, and Alan J. McKane. 2010. Evolutionary dynamics, intrinsic noise, and cycles of cooperation. *Phys. Rev. E* 81 (Jun 2010), 066122. Issue 6.

[4] Samuel Bowles and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and Its Evolution.* Princeton University Press, Princeton, NJ.

[5] Alessio Cardillo, Jesús Gómez-Gardeñes, Daniele Vilone, and Angel Sánchez. 2010. Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks. *New Journal of Physics* 12, 10 (Oct 2010), 103034.

[6] Jung-Kyoo Choi and Samuel Bowles. 2007. The coevolution of parochial altruism and war. *Science* 318, 5850 (2007), 636–640.

[7] David Easley and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World.* Cambridge University Press, Cambridge, UK.

[8] James H. Fowler and Nicholas A. Christakis. 2010. Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences* 107, 12 (2010), 5334–5338.

[9] Sergey Gavrilets. 2012. Collective action problem in heterogeneous groups. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, 1597 (2012), 1543–1556.

[10] Mark Granovetter. 1978. Threshold models of collective behavior. *Amer. J. Sociology* 83, 6 (1978), 1420–1443.

[11] Fritz Heider. 1946. Attitudes and cognitive organization. *The Journal of Psychology* 21, 1 (1946), 107–112.

[12] Josef Hofbauer, Karl Sigmund, et al. 1998. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, UK.

[13] Henrik Jeldtoft Jensen. 2022. *Complexity Science: The Study of Emergence.* Cambridge University Press, Cambridge, UK.

[14] Donella H. Meadows. 2008. *Thinking in Systems: A Primer.* Chelsea Green Publishing, White River Junction, VT, USA.

[15] John H. Miller and Scott E. Page. 2009. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life.* Princeton University Press, Princeton, NJ, USA.

[16] Melanie Mitchell. 2009. *Complexity: A Guided Tour.* Oxford University Press, Oxford, UK.

[17] Mark Newman. 2018. *Networks.* Oxford University Press, Oxford, UK.

[18] Martin A. Nowak. 2006. Five rules for the evolution of cooperation. *Science* 314, 5805 (2006), 1560–1563.

[19] Martin A. Nowak and Robert M. May. 1992. Evolutionary games and spatial chaos. *Nature* 359, 6398 (1992), 826–829.

[20] Elinor Ostrom. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge University Press, Cambridge, UK.

[21] Elinor Ostrom. 2000. Collective action and the evolution of social norms. *Journal of economic perspectives* 14, 3 (2000), 137–158.

[22] David G. Rand, Samuel Arbesman, and Nicholas A. Christakis. 2011. Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences* 108, 48 (2011), 19193–19198.

[23] David G. Rand, Alexander Peysakhovich, Gordon T. Kraft-Todd, George E. Newman, Owen Wurzbacher, Martin A. Nowak, and Joshua D. Greene. 2014. Social heuristics shape intuitive cooperation. *Nature Communications* 5, 1 (2014), 3677.

[24] Jonathan Skaggs and Jacob W. Crandall. 2026. Modeling Human Behavior in a Strategic Network Game with Complex Group Dynamics. In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems.* ACM, Pathos, Cyprus.

[25] Jonathan Skaggs, Michael Richards, Melissa Morris, Michael A. Goodrich, and Jacob W. Crandall. 2024. Fostering Collective Action in Complex Societies using Community-Based Agents. In *Proceedings of the International Joint Conference on Artificial Intelligence.* IJCAI, Jeju, South Korea, 211–219.

[26] Brian Skyrms. 2014. *Evolution of the Social Contract.* Cambridge University Press, Cambridge, UK.

[27] Qi Su, Benjamin Allen, and Joshua B Plotkin. 2022. Evolution of cooperation with asymmetric social interactions. *Proceedings of the National Academy of Sciences* 119, 1 (2022), e2113468118.

[28] György Szabó and Gábor Fáth. 2007. Evolutionary games on graphs. *Physics Reports* 446, 4 (2007), 97–216.

[29] Peter D. Taylor and Leo B. Jonker. 1978. Evolutionary stable strategies and game dynamics. *Mathematical biosciences* 40, 1-2 (1978), 145–156.

[30] Michael Tomasello. 2009. *Why We Cooperate.* MIT Press, Cambridge, MA.

[31] Duncan J. Watts. 2002. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences* 99, 9 (2002), 5766–5771.

[32] Richard Wrangham and Dale Peterson. 1996. *Demonic Males: Apes and the Origins of Human Violence.* Houghton Mifflin, Boston, MA.