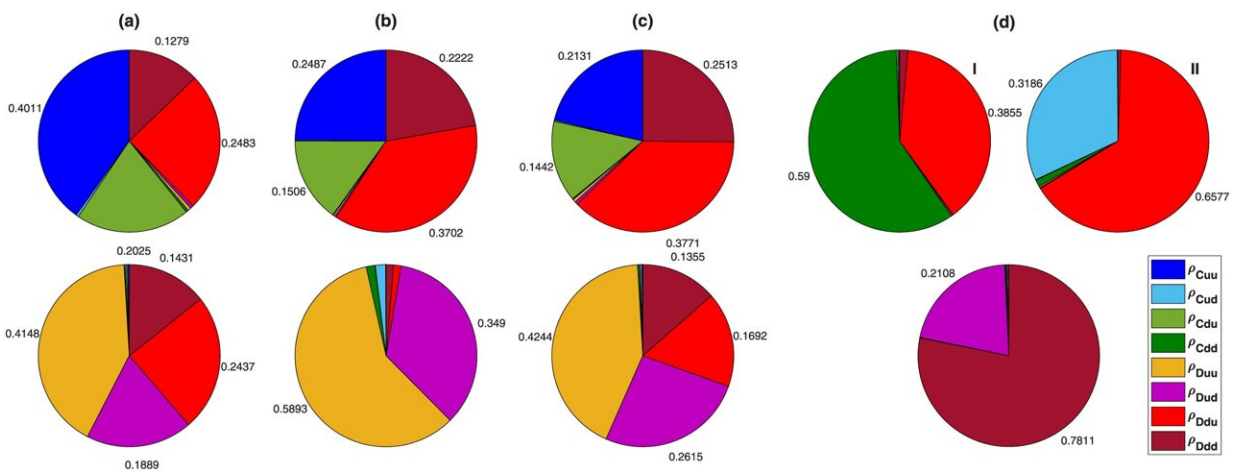# Understanding social dilemmas through games—the development of moral norms encourages selfless behavior

September 30 2022, by Jana Gregor

Frequency of different strategies in the fixed points of the replicator-mutator dynamics. Form A to D the game B is respectively, Snow Drift, Battle of the Sexes, Leader, and Stag hunt game. For anti-coordination games (A to C) the replicator-mutator dynamic has two stable fixed points, a cooperative fixed point where cooperation evolves (top) and a defective fixed point where cooperation does not evolve. For the Stag Hunt game, D, the replicator-mutator dynamic has two cooperative fixed points (top) and a defective fixed point (bottom). Credit: *PLOS Computational Biology* (2022). DOI: 10.1371/journal.pcbi.1010429

Human decision-making and the interplay of individual and group dynamics is incredibly complex. Unfortunately, our behavior can lead to

detrimental phenomena, such as the depletion of common sources.

Mohammad Salahshour, a researcher at the Max Planck Institute for Mathematics in the Sciences, has addressed the question of how individual strategic choices, social norms, and morality influence the decision-making process. His game-theory-based approach elucidates how the complexity of real-world strategic settings can lead to the evolution of moral norms, which help societies to better govern themselves by channeling individuals' decision-making in the interest of groups.

The process of decision-making can be conflict-ridden and can lead to social dilemmas that pit individual interests against benefits for the group or society. Morality provides a way out of this tragedy of the commons by promoting altruistic incentives and motivating individuals to curb selfishness and cooperate, even at a personal cost.

The emergence of morality is still an evolutionary puzzle. The central question is why a person should self-sacrifice and undermine their individual position to cooperate and benefit the group? It turns out that the individual pursuit of order and organization in society drives this development. So after initially striving for a form of social order out of pure self-interest, the ensuing moral system demands some form of self-sacrificing cooperation.

Individuals in groups often simultaneously face different strategic problems to be solved. Max Planck researcher Mohammad Salahshour employed elementary strategic games as a metaphor for a wide variety of these issues, including social dilemmas and coordination- and cooperation issues, such as the division of resources. In order to test whether these simple game theoretic approximations hold any merit in representing complex real-world interactions, he devised a novel evolutionary model of coupled interacting games.

In a first step, the individuals have to solve a Prisoner's Dilemma followed by a second game which can belong to different classes representing strategic scenarios individuals in groups may face. Studying the resulting Nash equilibria, Mohammad Salahshour could prove that the outcome of the players' choices in the social dilemma presented to them in the first game affect their strategic choices in the second game and can help solve different strategic problems, such as coordination, resource division, and leader choice.

Such an increased complexity of interacting games leads to a wide array of possible scenarios since a cooperating player might now be compensated for their subversion in the social dilemma. Depending on this compensation, the emergence of moral norms such as "good" or "bad" behavior occurs naturally: in the event of a low payout from the non-social dilemma game—and thus small coupling of the games and low complexity—there is no intrinsic value in cooperation and defection remains the rational choice.

However, as the coupling becomes strong enough, a symmetry-breaking phase transition appears, the symmetry between cooperation and defection breaks, and a set of cooperation-favoring social norms evolve, according to which cooperation stands as a valuable trait worthy of adoption.

Salashour's study of the evolution of moral norms revealed the existence of two quite distinct functions of morality. The previously mentioned promotion of self-sacrificing or altruistic behavior and the encouragement of mutually beneficial conduct. This second function does not imply self-sacrifice and could manifest in, for example, mutualistic cooperation or conflict resolution, norms that can promote social order and organization.

The mathematician says, "a moral system behaves like a Trojan horse:

once established out of the individuals' self-interest to promote order and organization, it also brings self-sacrificing cooperation and suppresses anti-social behavior." Interestingly his theory predicts that it is solely the cost of norms, not their benefit, that decides their establishment.

This fact can explain the surprising development of harmful social norms like destructive cultural practices, honor killings, or severe punishments. These norms are costly for the individual and often lack immediate social benefit resulting in a collective cost; however, they can be just as effective in promoting social order and stabilizing societies, especially in the absence of law-enforcement agencies.

The research was published in *PLOS Computational Biology*.

Provided by Max-Planck-Institut für Mathematik in den Naturwissenschaften