Diagnosing Social Dilemmas in Commons

Bryan Bruns[1]

bryanbruns@bryanbruns.com

Working Draft, February 8, 2021

Abstract: Changes in the ranking of possible outcomes map the diversity of collective action problems and solutions in elementary social situations, including the potential to "change the game" as in transforming a Prisoner's Dilemma into a Stag Hunt. Research on commons is often framed in terms of the need to solve social dilemmas, situations such as a tragedy of the commons where individual motives may conflict with achieving cooperation that could be better for everyone. Conceptualization of collective action in such situations is often based on a few simple models of symmetric games and their associated stories, namely Prisoner's Dilemma, Chicken, and Stag Hunt/Assurance. However, even in the payoff space of elementary social situations with two actors, two choices, and four ranked outcomes there are diverse problems for collective action. Most of the possible payoff structures do not involve social dilemmas, most situations are asymmetric, most yield unequal results at equilibrium, and most are only one or two steps away from transformation into win-win situations. Furthermore, information about outcomes is often incomplete. Outcomes and their valuation may vary with time and other factors. Diagnosing which kind of social situation may be present is important since different incentive structures pose different challenges for collective action and different opportunities for solutions, including transformations into situations with better solutions. This paper uses the Robinson-Goforth topology of payoff swaps in 2x2 games to map relationships among different types of collective action problems; identify a set of simple questions to distinguish between different problems even in the presence of limited information about outcomes; and discuss implications for diagnosis and potential solutions.

Introduction

Social dilemmas pose conflicts between individual incentives and achieving cooperation that would be mutually beneficial (Dawes 1980; Kollock 1998a; Van Lange et al. 2014). Different social situations pose different opportunities and risks for collective action, such as deterring defection from cooperation, building trust, or avoiding injustice. Payoff matrices for situations where two actors each have two choices offer elementary models of such problems. These are often named after stories that exemplify issues involved, such as Prisoner's Dilemma, Stag Hunt, and Chicken, (Luce and Raiffa 1957; Rapoport, Guyer, and Gordon 1976).

---

[1] Independent Researcher and Consultant, Greenbelt, Maryland, USA. bryanbruns@bryanbruns.com
An earlier version of this paper was presented at the Workshop on the Ostrom Workshop (WOW6) Conference Indiana University Bloomington, June 19–21, 2019.

As a simple example of potential cooperation, farmers in northeast Thailand would sometimes build small weirs to irrigate their crops (Bruns 1991). Two farmers on either side of a small stream could often do better by working together to build a shared structure that diverts water into their fields. However, the costs and benefits of a joint effort might differ. The possible incentive structures for collective action can take various forms (Taylor and Ward 1982). A joint effort might be necessary and best for both, while one person's efforts would be wasted, posing a Stag Hunt-type problem of how to assure coordination on the best outcome (Rousseau 2004; Hume 2003; Sen 1967; Runge 1986; Skyrms 2004). The effort might be better if both share the work, but each could be tempted to free ride, trying to get the other to do most of the work while the shirker reaps more of the benefits (Olson 1971), a situation which can be modeled as a Prisoner's Dilemma. Or, either might be capable of building a weir, after which the other could irrigate with little or no extra effort, but without a weir neither might get a crop. One who is willing to be more manipulative, aggressive, deceptive, or just less inclined to effort may be able to get the other to do most or all the work. This creates an incentive structure like the game of Chicken, also discussed as Hawk-Dove and Snowdrift (Kümmerli et al. 2007). These situations, and their multi-person analogues such as the Tragedy of the Commons, have been the focus of research on collective action to understand why cooperation might fail even if it could make everyone better off and to examine how cooperation might develop (Olson 1971; Hardin 1968; Axelrod 1984; Ostrom 1990; 2007; Nowak and Highfield 2011).

Analysis of social dilemmas has concentrated on conflicts between individual incentives and collective benefits, particularly the temptation to defect from cooperation or to avoid risk and thereby fail to cooperate in ways that would make both better off (Dawes 1980; Dawes and Messick 2000; Kollock 1998a; Van Lange et al. 2014). This paper applies the Robinson-Goforth topology of payoff swaps in 2x2 games map the diversity of elementary collective action situations and potential solutions and to show that:

1. In the space of possible 2x2 games, there are multiple, sometimes overlapping, kinds of collective action problems;
2. Most situations are not social dilemmas, most situations are asymmetric, most yield unequal outcomes at equilibrium, and most are only one or two steps away from transformation into situations with win-win solutions;
3. Even with limited information on payoffs, social dilemmas can be distinguished from other situations and identified in terms of the presence or absence of inefficient outcomes (Pareto-inferior), win-win outcomes, dominant strategies, or a worst outcome to avoid;
4. Uncertainty and variability in outcomes (trembling payoffs) increase the importance of diagnostics, principles, and problem-solving heuristics that can cope with diverse and dynamic situations, including potential transformations to "change the game" such as turning a Prisoner's Dilemma into a Stag Hunt.

Methods

Social dilemmas can be narrowly defined in terms of Prisoner's Dilemma situations where two dominant strategies (moves that are better regardless of what the other does) lead to a Pareto-inferior outcome, while there is a cooperative outcome that would be second-best for both and better than alternating between the other two outcomes. However, the term social dilemma, and related concepts such as the free-rider problem are often used more loosely to cover a variety of situations where individual incentives may conflict with cooperation that would be better for both.

The topology of payoff swaps in 2x2 games uses change in the ranking of outcomes to map the diversity of collective action problems in elementary situations. This allows both distinguishing between different situations and seeing their similarities. Transformation of Prisoner's Dilemma into a Stag Hunt by switching the ranking of the two top-ranked outcomes (3><4) offers a simple example, as shown in Figure 1. Switching the lowest two payoffs turns Prisoners' Dilemma into Chicken. Switching the top two payoffs turns Chicken into a win-win (no-conflict) game of Concord (Bruns 2018).

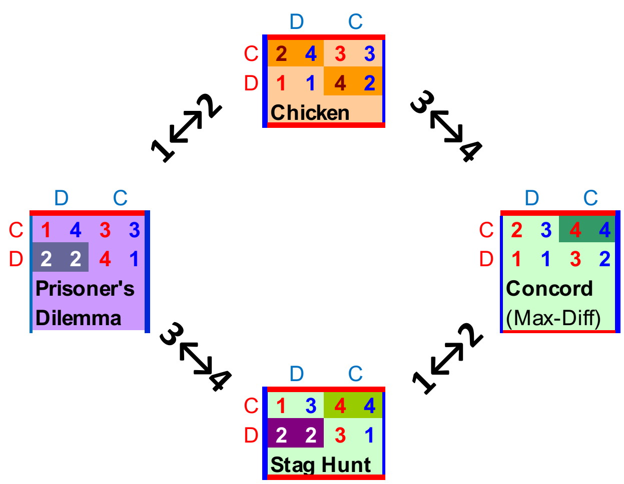

Figure 1 Payoff swaps link social dilemmas

Following the logic of how games are linked by payoff swaps, Robinson and Goforth found that the topology of connections between games could be displayed in a "Periodic Table" format as in Figure 2. Neighboring games are linked by changes in the ranking of outcomes. Twelve strict symmetric ordinal games make up a diagonal axis. Their payoffs combine to form asymmetric games. Pairs of games on either side of the axis of symmetric are equivalent by switching position as row or column player. As discussed in the results section, different kinds of games form distinct regions in the topology based on various properties including dominant strategies, the number of Nash Equilibria, and presence of Pareto-inferior equilibria. The topology shows relationships between different kinds of collective action problems, and can be used to clarify and systematize earlier classifications such as that by Holzinger (2003; 2008), as discussed below.

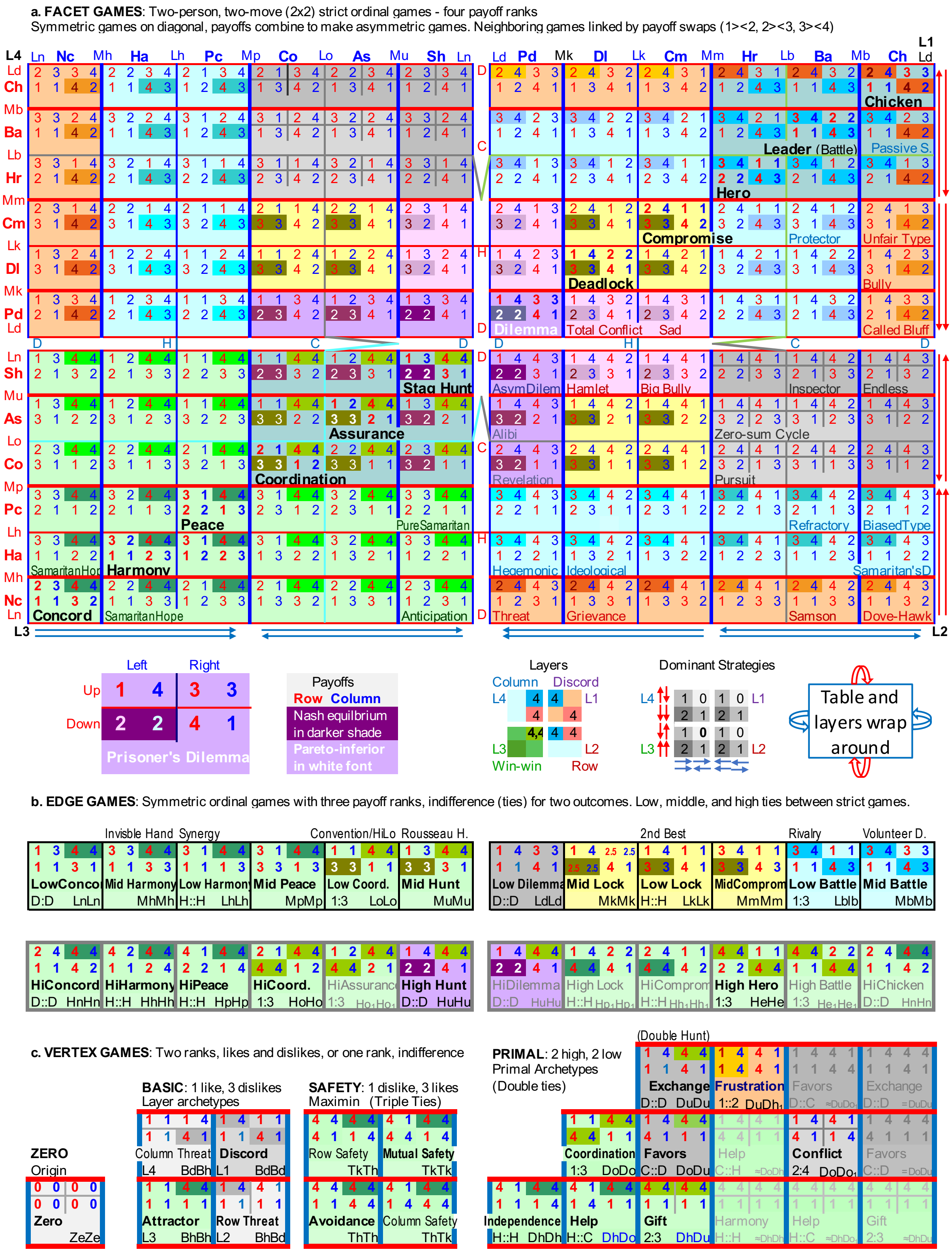

Figure 2 Periodic Table of Elementary Social Situations

Asymmetric games have been relatively neglected by researchers (Ernst 2005). However, asymmetric social dilemmas may occur, such as the one formed by combining payoffs from Stag Hunt and Prisoner's Dilemma. One actor has a dominant strategy, and the resulting choice for the other player leads to the Pareto-inferior outcome. In this case, cooperation would yield the best outcome for one and second-best for the other, in contrast to second-best (3,3) for both in the cooperative outcome for Prisoner's Dilemma. To explain how this might occur, Robinson and Goforth offers a variant of the Prisoner's Dilemma story where one of the two prisoners has an alibi. The topology helps show the extent of asymmetric games.

In the display, four "layers" differ by whether the best payoffs are in the same (win-win) outcome cell, as in Stag Hunt (*ShSh*) and Concord (*NcNc*)in the lower left Layer 3; diagonally opposite cells, as in Prisoners' Dilemma (*PdPd*) and Chicken (*ChCh*) in the upper right Layer 1; or the same row or column, as in Samaritan's Dilemma (*HaPd/PdHa*) and Cyclic Games in the upper left and lower right Layers 2 and 4.

Within each layer are four quadrants: in one quadrant both have dominant strategies leading to a single Nash equilibrium, a better move whatever the other does. In two adjoining quadrants, one or the other has a dominant strategy, which the second actor could anticipate and then make a move that is clearly best, again leading to a single equilibrium. In the other quadrant, there are (for strictly ordinal payoffs or pure strategies) either two equilibria, as in Stag Hunt and Chicken or else no equilibrium, as in cyclic games where one or the other would always prefer to move to a different outcome.

The topology of 2x2 games was originally developed for strict ordinal games, where the four possible outcomes can be strictly ranked by preference, as illustrated by payoffs from one to four. However, the topology can extend to include games with ties (indifference between outcomes) (Robinson, Goforth, and Cargill 2007). These can be visualized as lying between adjoining strict games, formed by making ties, "half-swaps." Situations with indifference between two or more outcomes can be presented by ties for payoff values.

Simplifying games by making ties provides a way to identifying archetypal models of social relationships, such as simple models for coordination, independence, and exchange (Bruns and Kimmich 2021). In the simplest situations, as shown at the bottom of Figure 1, there may be a single outcome that all wish to avoid, or an attractive win-win result, or conflict over outcomes, which might be resolved by taking turns. Simple models of games with ties can still help identify opportunities and challenges for cooperation, including aligned incentives, selection between alternative equilibria, potential exchange where each controls the other's outcome but has no control over their own outcome, and well as various forms of asymmetric power. For diagnosis, games with ties can also represent situations with incomplete information, where all that is known is that some outcomes rank better than others.

Payoff values can also be normalized and mapped onto the topology (Goforth and Robinson 2012; Bruns 2010; 2015). The topology maps the space of possible 2x2 games.

To the extent that payoffs are generated randomly, then payoff structures will tend to occur in the proportions shown in the topology (Simpson 2010). For actual situations, the frequency of different incentive structures is an empirical question, which may be related to resource characteristics, production functions for joint action, and other factors. However, unless or until more specific information is available about a situation, the proportions of games in the space of possible games as well as the frequency with a random distribution of payoffs offer one reasonable default expectation for discussing the possible prevalence of different situations.

Game theory analysis often makes a series of narrow and unrealistic assumptions including complete information, simultaneous moves, interacting only once, no communications, only caring about one's own outcomes, and not being able to make binding agreements. Relaxing these assumption provides a framework for exploring how different factors can influence behavior in strategic situations. The ways in which payoff swaps can turn one game into another offer a basis for considering potential transformations, including the impact of variability in payoffs, differences in perceptions, ignorance, and deliberate changes in games.

Results

Collective Action Problems. Analyzing the structure of the topology helps distinguish between different kinds of collective action problems. The topology of 2x2 shows groups of games with similar characteristics and problems for collective action. The taxonomy of 2x2 games developed by Rapoport, Guyer, and Gordon looked at the presence of Nash equilibria and dominant strategies. Holzinger (2003) classified 2x2 games in a table with cross-cutting categories. In one dimension these were based on the number of Nash equilibria and whether the equilibria were Pareto optimal. Holzinger's other dimension concerned the degree of conflict and equality or inequality in equilibrium outcomes. Games were then grouped in a typology of collective action problems. A refined version of Holzinger's typology can be mapped onto the topology of 2x2 games, as in Figure 3. This yields the following types of games, collective action problems, and potential solutions.

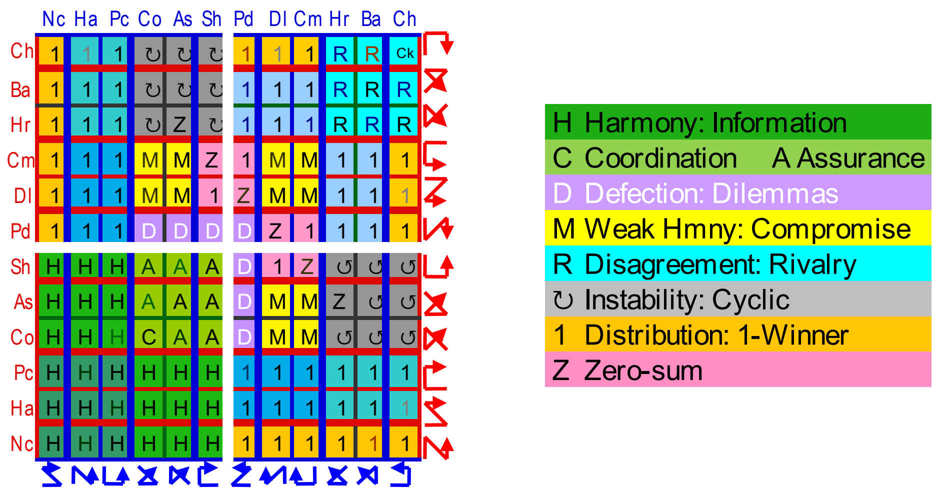

Figure 3 Types of Collective Action Problems

- *Ignorance and common knowledge in harmonious games*. Some games have a win-win equilibrium where both get their best outcome, which can result from dominant strategies. However, even when incentives are well-aligned, getting to win-win may depend on individuals have enough information to understand the potential outcomes of their choices. Often it would also depend on being able to anticipate the other's choices, some degree of common knowledge, as shown for games where only one actor has a dominant strategy, which make up half of the total of possible 2x2 games.
- *Inequality in one-winner games*. Most games have unequal outcomes where one winner gets best and the other does worse. In economic terms, these pose "distribution problems." These games with only one winner have been called suasion games, since the dissatisfied actor may try to find ways to achieve a different result (Martin 1992; Stein 1982). Holzinger used the term "Rambo" games, since one actor can get their best result without regard to what the other wants (Zürn 1993). These one-winner games include some with larger disparities in outcomes, 4,2 rather than 4,3 as well as games where neither gets their best outcome (3,2).
- *Inefficiency resulting from dominant strategies*. In addition to the symmetric Prisoner's Dilemma, there are also asymmetric dilemmas where a dominant strategy for one actor still leads to an equilibrium with payoffs inferior to another outcome that would be better for both.
- *Compromise settling for second-best.* Another set of games have an equilibrium with second-best results for both (3,3). Holzinger called these "weak harmony." They represent concepts such as compromise and "the best is the enemy of the good."

- *Instability in the absence of equilibrium.* Cyclic games without an equilibrium (in pure, unmixed, strategies) make up another distinct category. From any particular outcome, one or the other would always prefer to change their move. Some of these have no obvious stopping point to escape cycles of change, while others offer potential focal points for cooperation, which may be equitable or inequitable (4,3 or 3,3).
- *Insecurity and risk-avoidance*. In a choice between two alternative equilibria, one may be the safest choice for avoiding risk, while another equilibrium would be better for both. These have usually been described as stag hunt or assurance problems. In some asymmetric games of this type, only one actor has a payoff structure where risk conflicts with reaching the best outcome for both.
- *Rivalry over alternative equilibria*. Two possible equilibria may allow one to get their best outcome while the other does worse. However, stubborn pursuit of the best outcome could lead to both getting the worst or second-worst outcome. There have usually been discussed in terms of the story of "Battle of the Sexes" about coordinating a choice between alternative entertainment choices (Luce and Raiffa 1957; Rapoport 1967). Holzinger characterized these as "disagreement problems."

Further distinctions may cross-cut the categories above.

- *Social dilemmas as a region.* As Robinson and Goforth (2005) found, social dilemmas are neighbors in the topology. They form a compact connected region, including asymmetric dilemmas formed by combining payoffs from symmetric games. Prisoner's Dilemmas and Stag Hunts form a contiguous region of sixteen games, each of which has a Pareto-inferior equilibrium. Discussion of free-riding and incentives that tempt away from cooperation often also include the neighboring game of Chicken, which poses problems of defection and of avoiding a worst outcome, together with the problem of choice between rival equilibria that favor one or the other actor.
- *Sensitivity in social dilemmas*. As identified by Robinson and Goforth, the Prisoner's Dilemmas and their neighbors form the most diverse region within the topology. Even changes in the lowest two payoffs can change the number of equilibria and the payoffs at equilibrium. Switching the lowest two payoffs in the asymmetric "Alibi Game" (*ShPd*) adjoining Prisoner's Dilemma, can turn it into a cyclic game with no equilibrium, or can yield a game (Called Bluff, *PdCh*) with a single highly unequal (4,2) equilibrium outcome. Switching the top two payoffs for one person can create a Stag Hunt with two equilibria, where both could get their best payoffs but may instead get stuck at second-worst. Switching middle payoffs yields poor (3,2) results with no possibility of both doing better. Swapping both middle payoffs forms a game where both get second-best. Thus, dilemmas and their neighbors not only pose the problems of Pareto-inferior outcomes but are also highly sensitive to changes in payoffs, which can result in very different incentive structures.

- *Coordination problems*. Problems of choosing between alternative equilibria include Stag-Hunt style conflicts beween payoff maximization and risk avoidance, as well as rivalry between alternative equilibria where one or the other does better. It also includes a strict game of Safe Coordination, where risk avoidance leads to win-win, along with simpler coordination games with ties.
- *Remediability*. Figure 4 shows the number of swap steps that would be required to reach win-win. Most games are only a single swap away from win-win, sometimes through multiple pathways. Other games would take two steps. The most difficult would take three steps. In "sad" games, at equilibrium, nobody wins. These cluster along a line of zero-sum games linking the two clusters of cyclic games (zero rank-sum for ordinal games). They include Midlock (*MkMk*) the only symmetric zero-sum game, located between Prisoner's Dilemma and Deadlock.

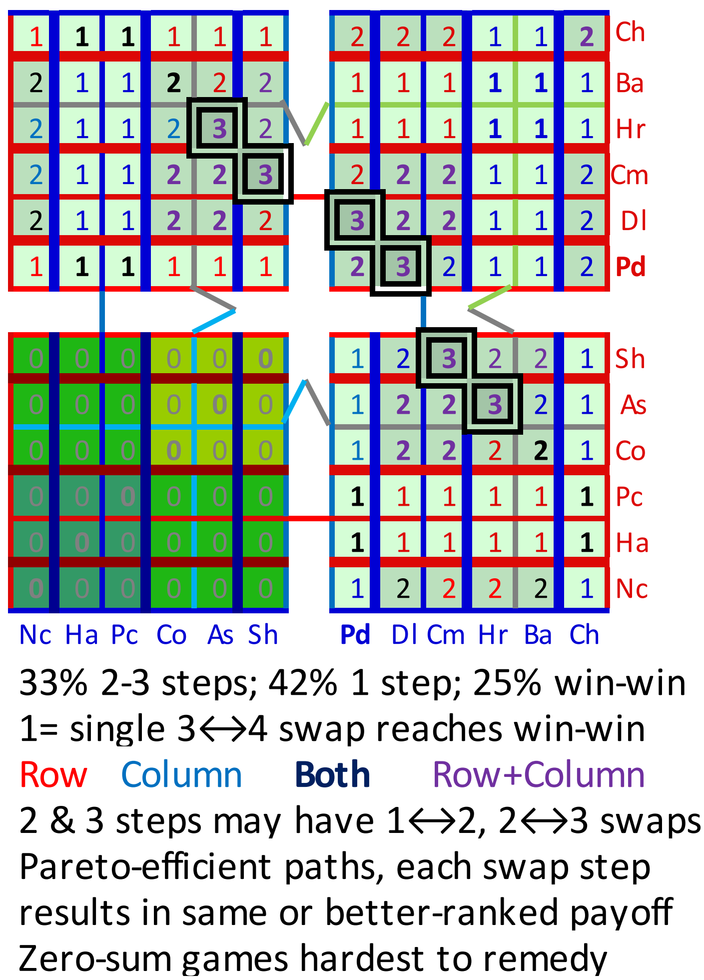

Figure 4 Payoff swap steps to reach win-win

Prevalence of social dilemmas and other problems. Most game theory research focuses on social dilemmas, particularly Prisoner's Dilemma, Chicken, and stag hunts (assurance problems). In terms of the space of possible games, as well as the likely frequency of different kinds of situations if payoffs occur randomly, then within the social dilemmas, Stag Hunts (8) are slightly more likely to occur than Prisoners' Dilemmas (7). (That count of stag hunts excludes Safe Coordination, CoCo, where risk avoidance does not conflict with payoff maximization for either actor.) As shown in Table1, other games would be far more likely to occur than the social dilemmas with Pareto-inferior equilibria.

Symmetric games are only one twelfth of the total. The vast majority of games are asymmetric. Most games do not have an equilibrium with equal payoffs. Games with equal payoffs at equilibrium are composed of the layer of win-win games, the family of second-best games, and three of the Prisoners' Dilemmas with equal payoffs at equilibrium, totaling slightly over a third of the games. Cyclic games have no equilibrium in pure strategies and make up another eighth of the total. Most games, a bit less than two-thirds, either yield unequal payoffs at equilibrium or lack an equilibrium. Thus, the preoccupation of research with a few symmetric social dilemmas does not match well with the proportions of what kinds of situations are possible, nor with the incentive structures most likely if payoffs occur randomly.

Table 1 Proportions of possible games

	Percentage	**Fraction**
Social Dilemmas	12%	17/144
Symmetric	8.3%	12/144
Equilibrium with equal payoffs	35.5%	51/144
• **Win-win (4,4)**	25%	36/144
• **Second best (3,3)**	8.3%	12/144
• **PD (2,2)**	2%	3/144
Cyclic-No equilibrium in pure strategies	12.5%	18/144
Equilibrium with unequal payoffs	52%	75/144

Discussion of social dilemmas typically assumes equality between players, as in symmetric games, and equality in outcomes, as with cooperative outcomes and Pareto-inferior equilibria in Prisoners' Dilemma and Stag Hunt, as well as in the cooperative outcome in Chicken. In terms of default expectations for randomly distributed payoffs, unequal payoffs at equilibrium, as in Chicken, would be much more likely to occur than pareto-inferior equilibria. Thus, in terms of a default distribution of games with randomly generated payoffs, inequality in outcomes could be a more prevalent problem than failure to achieve mutual gains through cooperation.

Diagnosing social dilemmas with incomplete information on payoffs. Some changes in payoffs may leave the outcomes and the basic structure and challenge for collective action unchanged or relatively similar. These show up as neighboring games in the visualization of the topology. This fits with Robinson and Goforth's assumption that swaps in the lowest two payoffs are likely to be the least significant changes. Looking at social dilemmas as a region of games connected by payoff swaps helps show how social dilemmas might be identified even if some outcome payoffs are uncertain or variable. For Stag Hunts, the key characteristic is whether there are two possible equilibria, one of which allows both to get their best outcome, while the other payoffs can occur in a variety of configurations. This occurs in the square block of nine games.

For the tragic incentive structure of Prisoner's Dilemma, the key characteristic is the presence of a dominant strategy leading to a Pareto-inferior Nash equilibrium, which occurs in the L-shaped set of seven games. The region with stag hunts and prisoners' dilemmas contains sixteen games with Pareto-inferior equilibria. Thus, social dilemmas can occur even when players face different incentive structures in asymmetric games or get unequal payoffs.

For analyzing social situations and identifying social dilemmas, there is no need to assume symmetry in payoff structures or equality in outcomes. Figure 5 provides a flow chart with diagnostic questions concerning whether there is a Pareto-inferior equilibrium and a win-win outcome, to distinguish assurance/stag hunt type problems from the tragic incentives of Prisoner's Dilemmas; or a second-best outcome from which both would like to defect but then would lead to the worst outcome for both, creating a Chicken-type situation. These key questions are sufficient to identify and distinguish social dilemmas, including those with asymmetric payoff structures and asymmetric Stag Hunts and Prisoner's Dilemmas where equilibrium payoffs are unequal. Figure 6 provides a more complete flow chart to diagnose the full set of possible payoff families for 2x2 games.

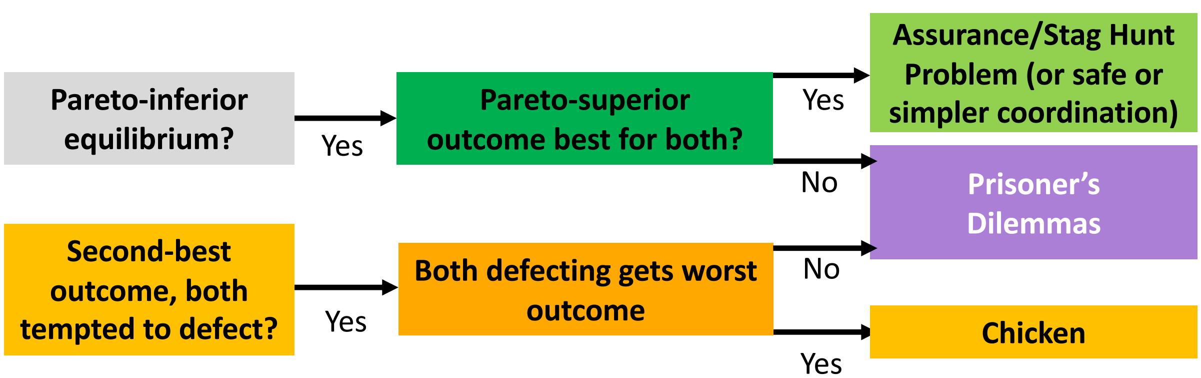

Figure 5 Diagnosing Social Dilemmas

Figure 6 Diagnosing Social Situations

Discussion

Coordination may be a more prevalent problem than tragic incentives. Most game theory research on social dilemmas has focused on Prisoner's Dilemmas, often on the particular payoff structure used by Axelrod (1984; Robinson and Goforth 2005). However, rather than a few exemplary games or specific payoff structures, social dilemmas can be better understood as a region of potential payoffs and incentive structures, characterized by the presence of equilibrium outcomes with Pareto-inferior payoffs. Furthermore, high swaps link Assurance and Safe Coordination, and their combinations with the rivalrous games of Hero and Battle into a larger region. Coordination problems, with two Nash equilibria may be much more frequent than tragic incentive structures where a dominant strategy leads to a single equilibrium.

Instability makes social dilemmas hard to identify and solve. In the topology, games where both get best or second best form a *sea of stability* where outcomes are relatively robust to changes in payoffs (darker green, blue, and yellow areas in Table 1) (Bruns 2015). However, in the social dilemma region, if there is uncertainty or instability (noisy or trembling payoffs) even for the lowest two payoffs, then it may be hard to be sure whether a social dilemma is present. This is particularly the case for the "wedge of chaos, formed by the symmetric and asymmetric Dilemmas. Potential changes in Prisoner's Dilemma may lead to other difficult situations, with poor payoffs, highly unequal payoffs, or no equilibrium, and thus different problems and opportunities for collective action. Stag Hunts are somewhat more robust to changes, and outcomes of payoff changes less diverse than Prisoners' Dilemmas. However, for assurance problems, changes in payoffs are still more prone to lead to neighboring games with more diverse equilibrium outcomes than games in other parts of the sea of stability.

Diagnosis. Incomplete or uncertain information about some specific outcomes does not necessarily prevent diagnosing social dilemmas, for which the presence of a Pareto-inferior outcome and an alternative where both could do better is a key characteristic. Furthermore, even with limited information, and uncertain or variable outcomes, the presence of a win-win equilibrium where both could get their best result helps distinguish assurance problems from the potentially tragic incentives for defection leading to a social trap in Prisoner's Dilemmas. Rather than simply saying there is a social dilemma, it is feasible and useful to identify and distinguish coordination and risk-avoidance problems where both could do best in contrast to the tragic incentives and inferior results of Prisoner's Dilemma-type situations. Similarly, incentives for "free riding" can exist in any of the social dilemmas, so more careful assessment is needed to understand the kind of "free rider" problem and possible solutions.

For repeated interaction in the Prisoners' Dilemma situation, cooperation can transform expected payoffs into a Stag Hunt structure (Skyrms 2004; Cole and Grossman 2010). Players then face a problem of assuring coordination. Depending on the actual payoffs, coordination could happen through choosing moves so that both get at least second-best, or by taking turns getting an outcome that allows a higher cumulative payoff. If play is repeated and payoffs can be measured in a comparable way, then a key question is whether taking turns getting unequal payoffs could yield a better total result compared to the "cooperative" move (Goforth and Robinson 2012). Games beyond a "recoordination line" where taking turns pays off better pose a somewhat different problem than those where repeated choice of the "cooperative" (second-best) outcome yields superior results. For repeated interaction, the potential to transform expected payoffs from a Prisoner's Dilemma to a Stag Hunt structure offers an additional reason why assurance problems may be more common and more important in the development of social order. However in practice, even in single-shot Prisoner's Dilemma situations, participants may perceive and act as if they were playing an assurance game (Kollock 1998a; 1998b).

Design principles and trembling payoffs. Uncertainty and variability in payoffs provide an additional reason why more general design principles may be superior to attempts to standardize on "one best way" panaceas (Ostrom 2007). Game theorists have analyzed "trembling hand" problems where people may make mistakes in their moves, intending to make one choice and actually doing something different (Selten 1975). Variability and uncertainty in payoffs can be thought about as a problem of "trembling payoffs." A typical explanation for Elinor Ostrom's development of more general principles for institutional design is the diversity of local conditions and the need for learning. If changes in payoff values may continually transform the collective action problems, this gives a further reason for not relying on a single type of solution or definition of the problem. For example, there may be seasonal shifts in the abundance of resources such as water, fish, pasture, and forest products. Ability to monitor others' behavior may shift with time of day, weather, terrain, season, and other factors. Violations that shift incentives towards cooperation may become harder to detect, and changes in social conditions could affect how likely sanctions are to be enforced. Confusion or dissension could affect expectations about how others may behave, undermining trust, reputation, and reciprocity that support cooperation. Conversely, improvements in common knowledge, consensus, and other factors may make conditions more favorable. Furthermore, such changes may differ between actors, creating asymmetric situations. Thus, a situation could move back and form between different kinds of collective action problems. This could occur in a deterministic and predictable way, as with seasonal rainfall and abundance of water, grass, forest products, etc. or cycles of animal reproduction. However, the underlying changes might be hard to predict or observe, creating uncertainty about the actual situation. Therefore, more general heuristics, solutions, and design principles may perform better than ones linked to a particular type of collective action problem. A map of the different kinds of overlapping problems of collective action can help understand the diversity of problems and potential solutions, and how uncertainty and variability contribute to the need to have multiple and flexible sets of solutions available, rather than a single standard remedy.

Dimensions of interdependence. Robinson and Goforth used Euler's Theorem to prove that their topology could be mapped onto a torus (i.e. a doughnut shape), if the torus had 37 holes. They also provide an practical visualization in their periodic table of 2x2 games in four layers, as discussed earlier, that shows most of the important relationships between 2x2 games. However, payoffs in 2x2 games can potentially vary within an eight dimensional space. Robinson and Goforth's original topology looked only at ordinal games. Normalized payoffs can be mapped onto a continuous version of the topology. However, this still fails to incorporate information about the relative importance of payoffs if they can be measured more precisely on interval or cardinal scales where the ranges of values may differ. Guisasola and Saari (2020) offer a way to decompose payoff matrices that includes a dimension for the incentives leading to Nash equilibria, but also dimensions for the ways in which potential gains from coordination and impacts of externalities can push behavior in other directions. Decompositions such as this offer a way to go beyond ordinal ranks and normalized payoffs to a fuller analysis of payoff space and factors that can influence cooperation.

Conclusions

The topology of 2x2 games provides a useful way to map and categorize different types of collective action problems. Different problems may require different solutions: common knowledge to converge on a win-win equilibrium, pragmatic willingness to compromise on second-best, trust to coordinate on the better of two alternative equilibria, mutual understanding of a focal point to escape from cyclic instability, reciprocity to take turns in repeated interaction, and caring about others and team spirit (other-regarding preferences and team reasoning) to overcome selfish temptations to defect from cooperation. To extent that payoffs are unknown or variable, a toolkit of problem-solving heuristics may be much more effective than relying on a single solution concept.

The Robinson-Goforth topology of payoff swaps in 2x2 games can be applied to identify social dilemmas not just in terms of a few exemplary games but as a region of similar social situations characterized by the presence of Pareto-inferior equilibria, in most of which the two players face different incentive structures and many of which have equilibrium payoffs that are not only Pareto-inferior but also unequal. Even with incomplete information, analysts can identify social dilemmas and distinguish coordination problems from situations with tragic incentives. Social dilemmas, particularly Prisoner's Dilemmas, are challenging not just due to the inefficiency of inferior outcomes, but also their sensitivity to changes in payoffs, making them harder to identify and to solve.

The prevalence of different kinds of collective action problems in commons is an empirical question, for which it is useful to have multiple models of different types of collective action. Some default expectations could be based on proportions of different models in the space of possible elementary payoff structures, as well as the likely frequency of games if payoffs occur randomly. Assuring coordination may be a more prevalent problem for governance than tragic temptations to defect from cooperation. Asymmetric problems of unequal opportunities and unequal outcomes are likely to be much more prevalent than inefficiency in achieving mutually beneficial (Pareto-superior) cooperation. Diagnosis of social situations in commons should pay more attention to distinguishing between different types of challenges for cooperation, including the difference between assuring coordination and deterring selfish defection, as well as the potential problems of and remedies for inequality in opportunities and outcomes.

References

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York, USA: Basic Books.

Bruns, Bryan. 1991. "The Stream the Tiger Leaped: A Study of Intervention and Innovation in Small Scale Irrigation Development in Northeast Thailand." Cornell University.

———. 2010. "Navigating the Topology of 2x2 Games: An Introductory Note on Payoff Families, Normalization, and Natural Order." *Arxiv Preprint ArXiv:1010.4727*.

———. 2015. "Names for Games: Locating 2 × 2 Games." *Games* 6 (4): 495–520. https://doi.org/10.3390/g6040495.

———. 2018. "From Tragedy to Win-Win Transforming Social Dilemmas in Commons." In . Georgetown Law Center, Washington DC, USA.

Bruns, Bryan, and Christian Kimmich. 2021. "Archetypes of Interdependence: Harmony, Coordination, and Exchange Generate Diverse Situations of Power, Cooperation, and Conflict."

Cole, Daniel H., and Peter Z. Grossman. 2010. "Institutions Matter! Why the Herder Problem Is Not a Prisoner's Dilemma." *Theory and Decision* 69 (2): 219–31.

Dawes, Robyn M. 1980. "Social Dilemmas." *Annual Review of Psychology* 31 (1): 169–93.

Dawes, Robyn M., and David M. Messick. 2000. "Social Dilemmas." *International Journal of Psychology* 35 (2): 111–16.

Ernst, Zachary. 2005. "A Plea for Asymmetric Games." *The Journal of Philosophy* 102 (3): 109–25.

Goforth, David, and David Robinson. 2012. "Effective Choice in All the Symmetric 2x2 Games." *Synthese* 187 (2): 579–605.

Guisasola, Santiago, and Donald Saari. 2020. "With Potential Games, Which Outcome Is Better?" *Games* 11 (3): 33.

Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science*, no. 162: 1243–48.

Holzinger, Katharina. 2003. "The Problems of Collective Action: A New Approach." Preprints aus der Max-Planck-Projektgruppe Recht der Gemeinschaftsgüter 2003/2. Bonn, Germany: Max-Planck-Projektgruppe Recht der Gemeinschaftsgüter.

———. 2008. *Transnational Common Goods: Strategic Constellations, Collective Action Problems, and Multi-Level Provision*. Springer.

Hume, David. 2003. *A Treatise of Human Nature*. https://www.gutenberg.org/ebooks/4705.

Kollock, Peter. 1998a. "Social Dilemmas: The Anatomy of Cooperation." *Annual Review of Sociology* 24 (1): 183–214.

———. 1998b. "Transforming Social Dilemmas: Group Identity and Cooperation." *Modeling Rational and Moral Agents*, 186–210.

Kümmerli, Rolf, Caroline Colliard, Nicolas Fiechter, Blaise Petitpierre, Flavien Russier, and Laurent Keller. 2007. "Human Cooperation in Social Dilemmas: Comparing the Snowdrift Game with the Prisoner's Dilemma." *Proceedings of the Royal Society B: Biological Sciences* 274 (1628): 2965–70.

Luce, Robert Duncan, and Howard Raiffa. 1957. *Games and Decisions: Introduction and Critical Survey*. New York, USA: Courier Dover Publications.

Martin, Lisa L. 1992. "Interests, Power, and Multilateralism." *International Organization* 46 (4): 765–92.

Nowak, Martin, and Roger Highfield. 2011. *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. Free Press.

Olson, Mancur. 1971. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

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

———. 2007. "A Diagnostic Approach for Going beyond Panaceas." *Proceedings of the National Academy of Sciences* 104 (39): 15181.

Rapoport, Anatol. 1967. “Exploiter, Leader, Hero, and Martyr: The Four Archetypes of the 2 Times 2 Game.” *Behavioral Science* 12 (2): 81.
Rapoport, Anatol, Melvin J. Guyer, and David G. Gordon. 1976. *The 2 x 2 Game*. Ann Arbor, USA: University of Michigan Press.
Robinson, David, and David Goforth. 2005. *The Topology of the 2x2 Games: A New Periodic Table*. London, UK: Routledge.
Robinson, David, David Goforth, and Matt Cargill. 2007. “Toward a Topological Treatment of the Non-Strictly Ordered 2x2 Games.” *Working Paper*.
Rousseau, Jean-Jacques. 2004. *A Discourse upon the Origin and the Foundation of the Inequality among Mankind*. Project Gutenberg. http://www.gutenberg.org/ebooks/11136.
Runge, Carlisle Ford. 1986. “Common Property and Collective Action in Economic Development.” *World Development* 14 (5): 623–35.
Selten, R. 1975. “A Reexamination of the Concept of Perfectness.” *International Journal of Game Theory*.
Sen, Amartya. 1967. “Isolation, Assurance and the Social Rate of Discount.” *The Quarterly Journal of Economics* 81 (1): 112–24.
Simpson, J. 2010. “Simulating Strategic Rationality.” Ph.D. Dissertation, Edmonton: University of Alberta.
Skyrms, Brian. 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge, UK: Cambridge University Press.
Stein, Arthur A. 1982. “Coordination and Collaboration: Regimes in an Anarchic World.” *International Organization* 36 (2): 299–324.
Taylor, M., and H. Ward. 1982. “Chickens, Whales, and Lumpy Goods: Alternative Models of Public Goods Provision.” *Political Studies* 30 (3): 350–70.
Van Lange, Paul AM, Daniel P. Balliet, Craig D. Parks, and Mark van Vugt. 2014. *Social Dilemmas: Understanding Human Cooperation*. Oxford, UK: Oxford University Press.
Zürn, Michael. 1993. “Problematic Social Situations and International Institutions: On the Use of Game Theory in International Politics.” In *International Relations and Pan-Europe: Theoretical Approaches and Empirical Findings; Publication of the Proceedings of the Inaugural Pan-European Conference in International Relations, Heidelberg, Germany, September 16-22, 1992*, 63–84. Hamburg, Germany: Verlag.