The Strategy Dynamics of Collective Systems

Ambrosio Valencia-Romero
The Roux Institute
Northeastern University
Portland, ME 04101, USA
a.valenciaromero@northeastern.edu

Paul T. Grogan
School of Computing and Augmented Intelligence
Arizona State University
Tempe, AZ 85287, USA
paul.grogan@asu.edu

I. Introduction

Designing systems-of-systems requires collective decision-making, which can be difficult due to local incentives, partial information, and decentralized control [1]. These complexities can make the process daunting, and conventional optimization methods provide effective solutions only sometimes [2]. Decentralized control leads to strategy-related uncertainty, affecting system performance and decision-making. Traditional engineering systems design relies on a central authority to decompose and allocate objective functions among stakeholders, which does not address the complexities of these settings. Game theory creates models and conducts analyses of strategic interactions between various actors to obtain insights into such situations and address related challenges [3–5]. However, more work is required to unlock the potential of game theory in these contexts. Addressing this research gap is crucial to accurately evaluate strategic uncertainty and its effect on the dynamics of collective action and comprehend how individuals make trade-offs in complex decision-making scenarios.

This work proposes a framework that generalizes the concepts of structural fear and greed strategy dynamics in two-strategy normal-form games, typically expressed for only two players, to any number of players. This approach builds on previous work that explored the impact of a design problem's strategic components on collective performance and aims to increase understanding of strategy dynamics across various multi-actor decision-making processes, regardless of their size or asymmetries. The concept of structural fear and greed space, which we call strategic hindrance, helps diagnose the stability of each player's incentives across all possible collective strategy outcomes and address strategy-related uncertainty. This approach focuses more on the players' beliefs about the game and the other players' strategies rather than their explicit knowledge. In this context, uncertainty arises when a player lacks 'common' or 'objective' knowledge about the likelihood of specific combinations of other players' strategies. This framework provides insights into the strategic trade-offs that decision-makers face when confronted with ambiguity and uncertainty about others' actions. By examining the landscape of possible strategic scenarios, individuals can better understand the various factors that come into play when making critical decisions in multi-actor environments.

II. Background and Conceptual Framework

The concept of fear and greed strategy dynamics is rooted in the study of social dilemmas through economic game theory. A "game" is an abstraction of a decision-making process involving several independent actors with diverse preferences over the decisions they could make and their possible consequences. The most common game representation used to describe social dilemmas is the *normal-form* (or *strategic-form*) game. A normal-form game $\mathcal{G}_{\mathcal{N}} = \left(\mathcal{N}, \left(\mathcal{S}_i\right)_{i\in\mathcal{N}}, \left(U_i\right)_{i\in\mathcal{N}}\right)$, consists of

- A player set $\mathcal{N}$ consisting of $n \geq 1$ player names or ids;
- A set of *pure* individual strategies $\mathcal{S}_i = \{s_i^{(1)}, \ldots, s_i^{(m_i)}\} \ni s_i$, with $m_i \geq 2$, for every $i \in \mathcal{N}$; and
- A *payoff* function $U_i : \mathcal{S} \mapsto \mathbb{R}^n$ valuing the preference of player i for each $s \in \mathcal{S}$.

An $n\times 2$ normal-form game is one in which $m_i = 2$ for every $i \in \mathcal{N}$, resulting in a total of 2^n pure collective strategies. Most classical social dilemmas are symmetric 2×2 in which the players share the same payoff function and strategy set. Let us set $\mathcal{S}_k = \{\varphi_k, \psi_k\}$ for every player $k \in \mathcal{N} = \{i, j\}$; in this work, we call $\varphi = \langle \varphi_1, \varphi_2, \ldots, \varphi_n\rangle$ the *status quo* collective strategy and $\psi = \langle \psi_1, \psi_2, \ldots, \psi_n\rangle$ the *desired outcome* from the perspective of a modeler, experimenter, or systems designer.

Classical 2×2 social dilemmas are commonly characterized by the differences in payoff between pursuing the same strategy and unilaterally deviating from it. The function

$$\ell_i(s_j) = \frac{1}{A_i}\Big[U_i(\varphi_i, s_j) - U_i(\psi_i, s_j)\Big] \in [-1, +1], \qquad (1)$$

where $A_i = \max_s U_i(s) - \min_s U_i(s) > 0$, is player i's *rescaled deviation loss* of deviating away from the status quo for a fixed s_j. The values

$$F_i(\varphi, \psi) = \ell_i(\varphi_j) \quad \text{and} \quad G_i(\varphi, \psi) = \ell_i(\psi_j)$$

are respectively known as the structural *fear* a structural *greed* factors [6]. The possible combinations of signed values of F_i and G_i characterize multiple 2×2 strategy dynamical domains exhibiting different payoff and risk dominance conditions [7, 8]. The following four regions stand out:

– *harmony* (HA):
$$F_i < 0 \quad \wedge \quad G_i < 0$$
– *coexistence* (CX):
$$F_i < 0 \quad \wedge \quad G_i > 0$$

$F_i < 0$	$G_i < 0$

sgn U_i		s_j	
		φ_j	ψ_j
s_i	φ_i	$-$	$-$
	ψ_i	$+$	$+$

e.g. concord game, peace game

(a) HA / Harmony

$F_i < 0$	$G_i > 0$

sgn U_i		s_j	
		φ_j	ψ_j
s_i	φ_i	$-$	$+$
	ψ_i	$+$	$-$

e.g. chicken, Bach or Stravinsky

(b) CX / Coexistence

$F_i > 0$	$G_i < 0$

sgn U_i		s_j	
		φ_j	ψ_j
s_i	φ_i	$+$	$-$
	ψ_i	$-$	$+$

e.g. stag hunt, coordination games

(c) BI / Bistability

$F_i > 0$	$G_i > 0$

sgn U_i		s_j	
		φ_j	ψ_j
s_i	φ_i	$+$	$+$
	ψ_i	$-$	$-$

e.g. prisoner's dilemma

(d) DE / Defection

Fig. 1. Main strategy dynamics in 2×2 social dilemma games in terms of the individual payoffs $U_i(s_i, s_j)$ — player ids i and j are interchangeable — and their relative differences ($+$ versus $-$). The labels "defect" and 'cooperate" commonly used in the literature are herein replaced by φ_i) and ψ_i, respectively. An economically rational player would presumably prefer those collective strategies marked with a plus sign ($+$) over those marked with a minus sign ($-$); for instance, a rational player in a harmony game would rather cooperate regardless of what they believe their partner would do, while players in a chicken game (coexistence dynamics) would need to anti-coordinate their strategies to reach a mutually beneficial outcome.

– *bistability* (BI):
$$F_i > 0 \quad \wedge \quad G_i < 0$$
– *defection* (DE):
$$F_i > 0 \quad \wedge \quad G_i > 0.$$

If the fear and greed values of both players coincide in one of the above regions, the game is guaranteed to have at least one collective strategy $s^\star \in \mathcal{S}$ from which no economically rational player would want to deviate, assuming the individual strategy of the other player is fixed. This is known as a pure-strategy Nash equilibrium (PNE) [9]; the set of PNE is $\mathcal{S}^\star \subseteq \mathcal{S}$. Figure 1 describes each of the four main strategy dynamics using a signed payoff matrix representation with every plus sign labeling the strategy that a rational player would prefer given the strategy of their partner. The sets $\mathcal{S}$ and $\mathcal{S}^\star$ in a 2×2 game are equal whenever

$$F_i = 0 \quad \wedge \quad G_i = 0,$$

for every $i \in \mathcal{N}$; we call this *indifference* dynamics (ZZ).

We generalize the structural fear and greed concept to characterize the strategy dynamics in more extensive and general collective decision-making processes ($n \geq 2$). This characterization aims to assess the stability of collective action in complex scenarios and assist in designing incentive mechanisms that mitigate unfavorable strategy dynamics and guide actors toward desired outcomes. The first step is to extend the definition of rescaled deviation losses, $\ell_i(s_j)$, to account for all conjectures that a single player can make about the actions that some or all of the other players in $\mathcal{N} \setminus \{i\} = -i$ could take. Let $\mathcal{K} \subset \mathcal{N} \setminus \{i\}$ be a proper subset of players and let $\mathcal{J} = \mathcal{N} \setminus (\mathcal{K} \cup \{i\})$ be the non-empty set of players different from i that are not in $\mathcal{K}$. Equation (1) can be generalized to

$$\ell_i(s_\mathcal{J}, s_\mathcal{K}) = A_i^{-1} \cdot \Big[U_i(\varphi_i, s_\mathcal{J}, s_\mathcal{K}) - U_i(\psi_i, s_\mathcal{J}, s_\mathcal{K}) \Big], \qquad (2)$$

with $s_{\text{-}i} = (s_\mathcal{J}, s_\mathcal{K})$, $s_\mathcal{J} \in \mathcal{S}_i^{|\mathcal{J}|}$, and $s_\mathcal{K} \in \mathcal{S}_i^{|\mathcal{K}|}$. We now consider the scenarios in which player i expects the players in $\mathcal{J}$ to coordinate their actions, assuming that the strategies of the players in $\mathcal{K}$ are fixed:

Definition 1. *Assume the players in $\mathcal{K}$ have adopted an immovable collective strategy $s_\mathcal{K} = \sigma_\mathcal{K}$. From the standpoint of player i, the normal-form game $\mathcal{G}_\mathcal{N}$ reduces to*

$$\mathcal{G}_{\mathcal{N}\setminus\mathcal{K}} = \Big(\mathcal{N}\setminus\mathcal{K}, (\mathcal{S}_i)_{i \in \mathcal{N}\setminus\mathcal{K}}, (U_i(\,\cdot\,, \sigma_\mathcal{K}))_{i \in \mathcal{N}\setminus\mathcal{K}} \Big). \qquad (3)$$

We call $\mathcal{G}_{\mathcal{N}\setminus\mathcal{K}}$ a player-reduced game observed by player i within $\mathcal{G}_\mathcal{N}$. Now, assume that player i conjectures that all players in $\mathcal{J}$ act as one player with individual strategy $s_\mathcal{J} \in \{\varphi_\mathcal{J}, \psi_\mathcal{J}\}$. The values of structural fear and greed specific to player i concerning an unanimous course of action by the players in $\mathcal{J}$ are given by

$$F_i^{(\sigma_\mathcal{K})}(\varphi, \psi) = \ell_i(\varphi_\mathcal{J}, \sigma_\mathcal{K}) \qquad (4)$$

and

$$G_i^{(\sigma_\mathcal{K})}(\varphi, \psi) = \ell_i(\psi_\mathcal{J}, \sigma_\mathcal{K}) \qquad (5)$$

The number of possible combinations of $\sigma_\mathcal{K}$ in (1) is given by

$$g_{(n)} = \sum_{\mathcal{K} \subset \mathcal{N}\setminus\{i\}} |\mathcal{P}(\mathcal{K})| = \sum_{|\mathcal{K}|=0}^{n-2} 2^{|\mathcal{K}|} \cdot \binom{n-1}{|\mathcal{K}|} = 3^{n-1} - 2^{n-1}, \qquad (6)$$

is the number of player-reduced games observed by player $i \in \mathcal{N}$ under the assumption of coordinating players in $\mathcal{J}$. For example, $g_{(3)} = 5$: in one instance, $|\mathcal{K}| = 0$ and $\mathcal{J} = \{j, k\} = -i$, thus φ_{-i} is restricted to $\{\langle \varphi_j, \varphi_k \rangle, \langle \psi_j, \psi_k \rangle\}$; in four other instances, $|\mathcal{K}| = 1$, with either $\mathcal{J} = \{j\} \Leftrightarrow \mathcal{K} = \{k\}$ or $\mathcal{J} = \{k\} \Leftrightarrow \mathcal{K} = \{j\}$, and the strategy of either player in $\mathcal{K}$ fixed to $\sigma_j \in \{\varphi_j, \psi_j\}$ or $\sigma_k \in \{\varphi_k, \psi_k\}$.

Definition 2. *The space of structural fear and greed values, or strategic hindrance, of player i in an n×2 game $\mathcal{G}_\mathcal{N}$ is*

$$\mathbf{H}_i^{(n)}(\varphi, \psi) = \begin{bmatrix} F_i^{(\sigma_\mathcal{K})}(\varphi, \psi) \\ G_i^{(\sigma_\mathcal{K})}(\varphi, \psi) \end{bmatrix}^T_{\times g_{(n)}} = \begin{bmatrix} \ell_i(\varphi_\mathcal{J}, \sigma_\mathcal{K}) \\ \ell_i(\psi_\mathcal{J}, \sigma_\mathcal{K}) \end{bmatrix}^T_{\times g_{(n)}}, \qquad (7)$$

where $g_{(n)}$ is the number of possible player-reduced games $\mathcal{G}_{\mathcal{N}\setminus\mathcal{K}}$ observed by i.

We visualize strategic hindrance spaces on the plane, with the structural fear along the vertical axis and the structural greed along the horizontal axis, and each quadrant named after each of the main strategy dynamics in Fig. 1. Figure 2 exemplifies how our approach is applied to characterize seven different 3×2 games.

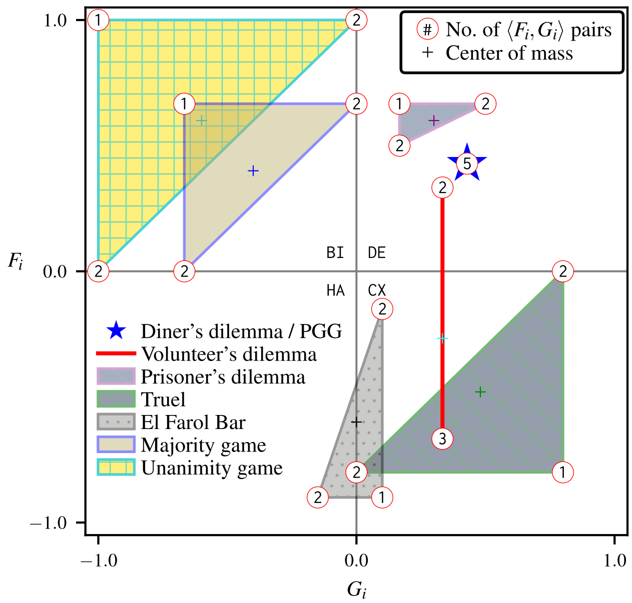

Fig. 2. Strategic hindrance spaces in seven classical 3×2 social dilemmas

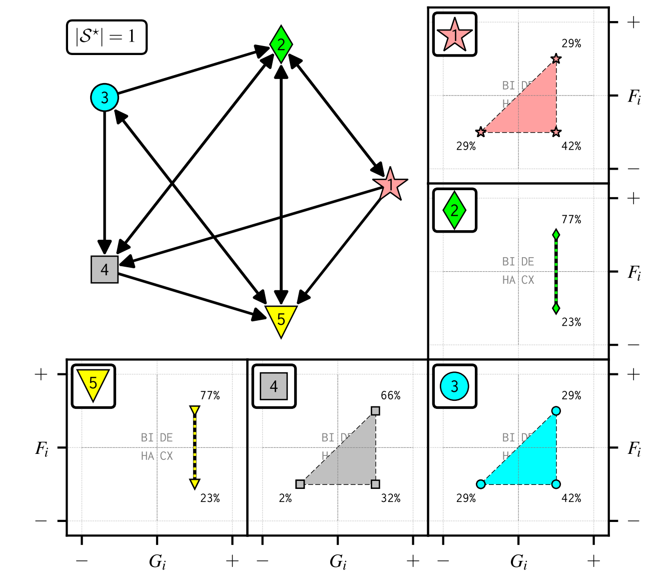

Fig. 3. Strategic hindrance in a 5-player graphical volunteer's dilemma. Payoff U_i follows Eq. (8), with $\mathcal{J} \cup \mathcal{K}$ as i's direct predecessors, and $a = b = 1$. The sole PNE is $\langle \psi_{\{1,3\}}, \varphi_{\{2,4,5\}} \rangle$, in contrast with the result for the normal-form (complete directed graphical) game, $\mathcal{S}^\star = \{\langle \psi_i, \varphi_{\text{-}i} \rangle \mid i \in \mathcal{N}\}$, $|\mathcal{S}^\star| = n$.

III. Application: Graphical Volunteer's Dilemma

In an n-player volunteer's dilemma, each player chooses between making a small contribution for the common good (ψ_i) or doing nothing and expecting to benefit from the contributions of others (φ_i). Player i's payoff in this game is defined as

$$U_i(s_i, s_{\mathcal{J}}, s_{\mathcal{K}}) = \begin{cases} 0 & \text{if } s_i = \psi_i \\ a & \text{if } s_i = \varphi_i \wedge s_{\mathcal{J}} = \psi_{\mathcal{J}} \\ -b & \text{otherwise,} \end{cases} \qquad (8)$$

where a and b are both positive real numbers. There are exactly n PNE in this game, each of them equal to $s^\star = \langle \psi_i, \varphi_{\text{-}i} \rangle$ for every $i \in \{1, 2, \ldots, n\}$ — every time one and only one player volunteers and the rest do not. Strategy ψ_i is *perfectly limited*, viz. it yields constant $U_i(\psi_i, s_{\text{-}i})$ for any $s_{\text{-}i} \in \mathcal{S}_i^{n-1}$, restricting interactive effects [4]. Figure 2 shows the below individual strategic hindrance space in a 3-player volunteer's dilemma with $a = 1$ and $b = 2$:

$$\mathbf{H}_i^{(3)}(\varphi, \psi) = \frac{1}{3} \begin{bmatrix} -2 & +1 & -2 & +1 & -2 \\ +1 & +1 & +1 & +1 & +1 \end{bmatrix}^{\mathrm{T}}.$$

Three structural fear and greed pairs fall on the coexistence region and two fall on the defection region. It can be shown that the greater the number of players in the volunteer's dilemma game, the greater the incidence of defection dynamics on each player's strategic hindrance, which could lead to an increased likelihood of no players volunteering at all. This quantitative insight into the evolution of defection dynamics in the volunteer's dilemma game is reminiscent of the prevalence of the diffusion of responsibility and bystander effect phenomena observed in real-world settings [10].

The interactions between players in a normal-form game can be represented as a complete directed graph. Games described using directed graphs are referred to as *graphical games*. In these games, player i's payoff is only affected by the strategies of the players that are their direct predecessors, i.e. there is a directed edge (arc) coming from i's neighbors.

Figure 3 shows our approach applied to a 5-player graphical volunteer's dilemma game. Harmony dynamics emerge for players 1, 3, and 4, with reduced connectivity resulting in lower levels of greed. The single PNE would see players 1 and 3 — the less connected — more inclined to volunteer (ψ_1 and ψ_3), while the rest do nothing (φ_2, φ_4, and φ_5).

IV. Discussion and Conclusions

While strategy dynamics for two-player two-strategy games neatly fall within four domains on the strategic hindrance plane of fear and greed, normal-form games with a greater number of players exhibit more complex dynamics. The strategy dynamics experienced by each player depend on the belief of other players' strategies and connectivity.

This framework shows how considering player-reduced games allows a similar mapping onto the strategic hindrance spaces with corresponding interpretations of the prevalence of different dynamics. A graphical method illustrates the strategy dynamics as a polygon region, with each node representing dynamics at granular, sub-game levels. The example application case shows that a volunteer's dilemma gravitates towards defection dynamics with increasing numbers of players — consistent with real-world observations of the bystander effect. It also shows that defection can be partially mitigated by aptly removing connections between players.

Applications of this framework can be used to understand complex interdependent dynamics among actors designing system-of-systems. Furthermore, based on this understanding, system-of-systems engineers can devise or develop incentive mechanisms to mediate unfavorable dynamics and stabilize socially efficient outcomes.

Acknowledgment

This material is based upon work supported in part by the National Science Foundation under Grant No. 1943433.

References

[1] M. W. Maier, "Architecting Principles for Systems-of-systems," *Systems Eng.*, vol. 1, no. 4, pp. 267–284, 1998.

[2] A. Gurnani and K. Lewis, "Collaborative, Decentralized Engineering Design at the Edge of Rationality," *ASME J. Mech. Des.*, vol. 130, no. 12, p. 121101, 2008.

[3] Z. Sha, K. N. Kannan, and J. H. Panchal, "Behavioral Experimentation and Game Theory in Engineering Systems Design," *ASME J. Mech. Des.*, vol. 137, no. 5, p. 051405, 2015.

[4] P. T. Grogan, K. Ho, A. Golkar, and O. L. de Weck, "Multi-actor Value Modeling for Federated Systems," *IEEE Syst. J.*, vol. 12, no. 2, pp. 1193–1202, 2018.

[5] P. T. Grogan, "Stag Hunt as an Analogy for Systems-of-systems Engineering," *Procedia Comput. Sci.*, vol. 153, pp. 177–184, 2019.

[6] T.-K. Ahn, E. Ostrom, D. Schmidt, R. Shupp, and J. Walker, "Cooperation in PD Games: Fear, Greed, and History of Play," *Public Choice*, vol. 106, no. 1-2, pp. 137–155, 2001.

[7] M. A. Nowak and R. M. May, "Evolutionary Games and Spatial Chaos," *Nature*, vol. 359, no. 6398, pp. 826–829, 1992.

[8] C. Hauert, "Effects of Space in 2×2 Games," *Int. J. Bifurcat. Chaos*, vol. 12, no. 07, pp. 1531–1548, 2002.

[9] J. F. Nash, "The Bargaining Problem," *Econometrica*, pp. 155–162, 1950.

[10] P. Campos-Mercade, "The Volunteer's Dilemma Explains the Bystander Effect," *J. Econ. Behav. Organ.*, vol. 186, pp. 646–661, 2021.

Ambrosio Valencia-Romero (he/him) received his B.S. degree in mechanical engineering from the Universidad del Atlántico in Barranquilla, Colombia, 2012; his M.S. degree in mechanical engineering from the Universidad de Puerto Rico at Mayagüez in 2016; and his Ph.D. degree in systems engineering from the School of Systems and Enterprises at Stevens Institute of Technology, in Hoboken, NJ, USA, in 2021. He is currently a Postdoctoral Research Fellow with the Roux Institute at Northeastern University in Portland, Maine. His research interests include strategic engineering design, agent-based and multi-agent systems modeling, and collective design experiments for the study of complex socio-technical systems.

Paul T. Grogan (he/him) is an Associate Professor with the School of Computing and Augmented Intelligence at Arizona State University in Tempe, Arizona. He leads the Collective Design Lab, which develops and studies the use of information-based methods and tools for engineering design in systems with distributed architectures such as aerospace, defense, and critical infrastructure. He holds a Ph.D. in engineering systems and a S.M. degree in aeronautics and astronautics from the Massachusetts Institute of Technology and a B.S. degree in engineering mechanics from the University of Wisconsin Madison.