Proceedings of the ASME 2018 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2018
August 26-29, 2018, Québec City, Québec, Canada

DETC2018-85786

TOWARD A MODEL-BASED EXPERIMENTAL APPROACH TO ASSESSING COLLECTIVE SYSTEMS DESIGN

Ambrosio Valencia-Romero*
School of Systems and Enterprises
Stevens Institute of Technology
Hoboken, NJ 07030
Email: avalenci@stevens.edu

Paul T. Grogan
School of Systems and Enterprises
Stevens Institute of Technology
Hoboken, NJ 07030
Email: pgrogan@stevens.edu

ABSTRACT

This work presents a conceptual model of collective decision-making processes in engineering systems design to understand the tradeoffs, risks, and dynamics between autonomous but interacting design actors. The proposed approach combines value-driven design, game theory, and simulation experimentation to study how technical and social factors of a design decision-making process facilitate or inhibit collective action. The collective systems design model considers two levels of decision-making: 1) lower-level design value exploration; and 2) upper-level design strategy selection. At the first level, the actors concurrently explore two strategy-specific value spaces with coupled design decision variables. Each collective decision is mapped to an individual scalar measure of preference (design value) that each actor seeks to maximize. At the second level, each of the actor's design values from the two lower-level design exploration tasks is assigned to one diagonal entry of a normal-form game, with off-diagonal elements calculated in function of the "sucker's" and "temptation-to-defect" payoffs in a classical strategy game scenario. The model helps generate synthetic design problems with specific strategy dynamics between autonomous actors. Results from a preliminary multi-agent simulation study assess the validity of proposed design spaces and generate hypotheses for subsequent studies using human subjects.

*Address all correspondence to this author.

1 INTRODUCTION

Design of engineering systems increasingly considers federated or decentralized architectures embodied in the artifact (e.g. grid energy storage or spacecraft constellations), the design team (e.g. joint projects across agencies), or supporting systems such as manufacturing (e.g. the Boeing 787). Distributed features on any of these dimensions offer a potential to improve performance but also diverge from a centralized engineering decision-making process [1]. Design literature emphasizes methods building on utility theory to aggregate preferences and decision theory to reach optimal outcomes [2, 3]. However, single- and multi-objective optimization methods alone do not reflect the local objectives, partial information, and decentralized control present in federated systems. In contrast, adjacent fields such as software, artificial intelligence, and robotics recognize issues of decentralization, producing a rich literature in multi-agent systems focused on operational decision-making and control [4].

Collective systems design—hereinafter referred to as CoDe—advances integrative theory that considers both the technical factors of a design problem and social interactions among entities undertaking the design activity [5]. Practical approaches to CoDe emphasize participation, mediation, and negotiation leveraging computer-based models [6–8]. Rather than prescribing design solutions, this paper takes a more theoretically-grounded approach by modeling the design process to explain outcomes based on underlying strategy dynamics defined using concepts from game theory. Improved understanding of the relationship between the fundamental problem structure and natural outcomes will help create better mechanisms and incentives to guide multi-actor design problems.

Game theory logically extends decision theory to multi-actor cases and has been explored in past design literature over several decades (e.g. see discussion in [9] and application cases in [10, 11]). Typically, a design alternative represents a game-theoretic strategy and corresponding Nash equilibria represent stable decision sets for non-cooperative games (which need not be optimal in terms of Pareto efficiency) [12]. This framing becomes somewhat problematic in practice because strategy analysis often takes place at a higher level of abstraction than most design decisions, necessitating a differentiation between strategy and technical design decisions [13]. Prior work frames design as a bi-level problem and studies scenarios with strategy dynamics of a Stag Hunt game where two Nash equilibria force design actors to balance preference for risk and reward [13]. Other strategy game scenarios remain generally unexplored.

The objectives of this paper are to 1) formulate a conceptual modeling framework for CoDe and 2) generate design decision-making problems with specified strategy dynamics. Outcomes from this paper will be used as synthetic design problems in future behavioral experiments. The proposed framework builds on the bi-level model in [13] to define two design phases: 1) design space exploration; and 2) strategy selection. First, actors explore strategy-specific design spaces to identify (individual) value-maximizing design alternatives. Second, actors choose among designs presented as pure strategy payoffs in a normal-form game characterized as one of four strategy dynamical domains (i.e. cooperation, coexistence, bistability, and defection) based on the "sucker's" and "temptation-to-defect" mixed-strategy payoffs [14]. Operated in reverse, fixing the "sucker's" and "temptation-to-defect" parameters allows generation of design problems with specified strategy dynamics.

Section 2 describes the model-based approach to CoDe research in more detail. Section 3 analyzes the results from a preliminary multi-agent simulation study of the proposed experiment. Section 4 provides a general discussion on the outcomes from the simulation study and the implications of the proposed approach in CoDe research. Finally, Section 5 provides some concluding remarks and opportunities for future research on collective decision-making processes.

2 COLLECTIVE SYSTEMS DESIGN MODEL

This section introduces and implements a model of CoDe as a generic design activity depicting individual and collective decisions in an engineering systems design process. The approach combines concepts from value-driven design [2,15,16] and game theory to study how technical and social factors of engineering systems design facilitate or inhibit collective action. The objective of the model is to represent the tradeoffs, risks, and dynamics between interacting design actors at different levels of collective effort; it does not attempt to prescribe "best" or "optimal" solutions to multi-actor design problems. Specifically, this model generates synthetic design problems with specified strategy dynamics between autonomous actors.

The following sections introduce the CoDe model as a bi-level decision-making problem with lower-level technical design decisions and upper-level strategy tradeoffs. Parameterized value functions for the two levels of design formulate synthetic design problems conforming to desired canonical two-strategy games.

2.1 Model Overview

The CoDe model is based on a bi-level decision-making problem with two interrelated phases: 1) design space exploration, and 2) strategy selection. Consider vector $s = [s_1, \dots, s_n]$ whose elements specify the strategy decision for each of the $n \geq 2$ actors from the symmetric strategy space $\mathbb{S} \times \cdots \times \mathbb{S} = \mathbb{S}^n$. The first part of the model illustrates lower-level technical design processes within the context of all actors under the same design strategy, i.e. diagonal strategy vector $\bar{s} = \mathcal{S}\vec{\mathbf{1}}$, $\mathcal{S} \in \mathbb{S}$. A design vector $d = [d_1, \dots, d_n]$ specifies design decisions for each of n actors from the design space $\mathbb{d} = \mathbb{d}_1 \times \cdots \times \mathbb{d}_n$ under $\bar{s}$. A multi-actor design value function $V^s(d) = [V_1^s(d), \dots, V_n^s(d)]$ maps the design d to a scalar measure of preference for each actor i, who then searches for value-maximizing designs $d_i^s \in \{d_i^{\bar{s}}\}$ under each $\bar{s} \in \mathbb{S}^n$ in Eqn. (1),

$$d_i^s \in \left\{ \arg \max_{d_i \in \mathbb{d}_i,\, \bar{s} \in \mathbb{S}^n} V_i^{\bar{s}}(d_i, d_{-i}) \right\}. \tag{1}$$

Each element of d_{-i} represents the design decision made by every other actor $j \in \{1, \dots, n\} \setminus \{i\}$.

While framed simply, interactive effects and conflicting local objectives contribute challenges to finding joint design decisions. More generally, Eqn. (1) only describes a locally optimal design decision for actor i given d_{-i} under diagonal strategy vector $\bar{s}$, but not under $s \neq \bar{s} \in \mathbb{S}^n$. Any deviation from $\bar{s}$ results in an upper-level decision-making process in which autonomous actors face different organizational and strategy tradeoffs. Using the multi-actor design value function $V^s(d)$ constrained to strategy-specific designs $d^s = [d_1^s, \dots, d_n^s]$ found in Eqn. (1), actors search for the payoff-maximizing strategy in Eqn. (2):

$$s_i^\star = \arg \max_{s_i \in \mathbb{S}} V_i^s(d_i^s, d_{-i}^s), \tag{2}$$

For $n = 2$ actors and $\mathbb{S} = \{\mathcal{R}, \mathcal{B}\}$, strategies **Red** and **Blue**, the resulting upper-level problem can be formulated as a normal-form game shown in Table 1. Generally, actors iterate through and be-

TABLE 1. DESIGN STRATEGY GAME FOR $n=2$ PLAYERS

		Player 2	
		Red ($\mathcal{R}$)	**Blue** ($\mathcal{B}$)
Player 1	**Red** ($\mathcal{R}$)	$V_1\left(d_1^{\mathcal{R}}, d_2^{\mathcal{R}}\right)$ $V_2\left(d_2^{\mathcal{R}}, d_1^{\mathcal{R}}\right)$	$V_1\left(d_1^{\mathcal{R}}, d_2^{\mathcal{B}}\right)$ $V_2\left(d_2^{\mathcal{R}}, d_1^{\mathcal{B}}\right)$
	Blue ($\mathcal{B}$)	$V_1\left(d_1^{\mathcal{B}}, d_2^{\mathcal{R}}\right)$ $V_2\left(d_2^{\mathcal{B}}, d_1^{\mathcal{R}}\right)$	$V_1\left(d_1^{\mathcal{B}}, d_2^{\mathcal{B}}\right)$ $V_2\left(d_2^{\mathcal{B}}, d_1^{\mathcal{B}}\right)$

tween both the design space exploration problem, Eqn. (1), and the strategy selection problem, Eqn. (2), in a time-constrained fashion. The next sections expand on the implementation of each part of the CoDe model for $n=2$ actors and $\mathbb{S}=\{\mathcal{R},\mathcal{B}\}$.

2.2 Design Model Implementation

This design model implementation generates synthetic design problems with two autonomous actors, i and j, within a strategy scenario $\bar{s}=\langle s_i, s_j\rangle \,|\, s_i=s_j=\mathcal{S}\in\{\mathcal{R},\mathcal{B}\}$. Each design space has a finite number of design variables $\mathbb{d}_i\in\{0,1,\ldots,K\}$ yielding $|\mathbb{d}_i|=K+1$ design alternatives for each actor i. Function $V_i^{\bar{s}}(d)\in V^s(d)$ is given by:

$$V_i^{\bar{s}}(d)=v_i\left(f^{\bar{s}}(d)\right), \tag{3}$$

where $f^{\bar{s}}(d)=\langle f_i^{\bar{s}}, f_j^{\bar{s}}\rangle$ is the strategy-specific technical design value function and v_i is a value-sharing mechanism through which actor i's expected reward is computed. The inclusion of function v_i in Eqn. (3) attempts to further explore how a value sharing policy in lower-level design decision-making processes impacts actors' behavior in upper-level strategy tradeoffs.

Design Value Function. The discrete value function in Eqn. (4) generates synthetic strategy-specific technical design value spaces for each symmetric actor i:

$$f^{\bar{s}}(d)=f^{\bar{s}}(d_i,d_j)=\left\lfloor \begin{array}{l} \dfrac{K}{2}e^{-\beta\left[(d_i-d_i^\star)^2+(d_j-d_j^\star)^2\right]} \\ +\displaystyle\sum_{u=0}^{K}\sum_{v=0}^{K}\varepsilon_{uv}e^{-\alpha\left[(d_i-u)^2+(d_j-v)^2\right]} \end{array}\right\rfloor_{\mathbb{V}} ; \tag{4}$$

with parameters $\alpha,\beta>0$, random sample $\varepsilon_{uv}\sim\mathcal{U}(0,1)$, function $\lfloor\ldots\rfloor_{\mathbb{V}}$ to map to a discrete value scale $\mathbb{V}=\{0,5,10,...,100\}$, and unique global maximum located at $d^\star=\langle d_i^\star, d_j^\star\rangle$. Equation (4) uses components from the first of two test optimization functions presented in [17]. The resulting technical value space, as observed by actor i, is expressed as a matrix $\mathbf{F}_{ij}^{\bar{s}}:\mathbb{d}\mapsto\mathbb{V}$ mapping the collective design space to the value scale. For a two-actor task, the value spaces are presented in a square grid where each cell shows value $f^{\bar{s}}(d_i,d_j)$ mapped to a perceptually uniform discrete color scheme (monotonically increasing lightness), Fig. 1.

An iterative sampling method generates the technical value spaces $f^{\bar{s}}(d)=f^{\bar{s}}(d_i,d_j)=\langle f_i^{\bar{s}}, f_j^{\bar{s}}\rangle$ to satisfy four constraints:

$$d_i{}^\star \neq d_j{}^\star, \tag{5}$$

$$d_{\text{sym}}^{\bar{s}}=\arg\max_{(d_i,d_j)\in\mathbb{d}|d_i=d_j} f^{\bar{s}}(d_i,d_j), \tag{6}$$

$$f^{\bar{s}}(d_{\text{sym}}^{\bar{s}})=f_{\text{sym}}^{\bar{s}}, \tag{7}$$

$$f_i^{\bar{s}}(d)+f_j^{\bar{s}}(d)<2\cdot f_{\text{sym}}^{\bar{s}}\;\forall\, d\neq d_{\text{sym}}^{\bar{s}}. \tag{8}$$

Equation (5) ensures the unique global maximum is not located on the plane of symmetry between the two actors, enforcing a tradeoff between individual and collective objectives. Equation (6) identifies a unique symmetric design decision maximizing individual value as the goal of a collective design. Equation (7) requires the corresponding value for the symmetric collective design to be a specified value $f_{\text{sym}}^{\bar{s}}$. Finally, Eqn. (8) requires the symmetric goal $d_{\text{sym}}^{\bar{s}}=\langle d_i^{\bar{s}}, d_j^{\bar{s}}\rangle$ to have larger value sum than alternatives.

In the CoDe model, design actors explore the technical value spaces associated with each strategy in $\mathbb{S}=\{\mathcal{R},\mathcal{B}\}$ and select design decisions $d^{\mathcal{R}}=\langle d_i^{\mathcal{R}}, d_j^{\mathcal{R}}\rangle$ and $d^{\mathcal{B}}=\langle d_i^{\mathcal{B}}, d_j^{\mathcal{B}}\rangle$ within a specified time period τ. Technical value spaces generated for particular tasks consider $K=6$ such that $|\mathbb{d}_i|=7$ and $|\mathbb{d}|=49$. The value space corresponding to $\mathcal{R}$ in Fig. 1 uses parameters $\alpha=1.00$ and $\beta=0.75$ with $f_{\text{sym}}^{\mathcal{R}}=75$ and $f_{\text{sym}}^{\mathcal{B}}=55$ (not shown).

Reward Mechanism Functions. A reward mechanism influences the effective value of each alternative during design space exploration. A general value-sharing mechanism, v_i, encourages actors to select collectively efficient outcomes by sharing $(1-r)\cdot 100\%$ of each other's value with one another:

$$v_i\left(f^{\bar{s}}(d)\right)=r\cdot f_i^{\bar{s}}+(1-r)\cdot f_j^{\bar{s}}, \qquad 0.50\leq r\leq 1.00 \tag{9}$$

Different values of r in Eqn. (9) are expected to have considerable impact on actors' behavior in upper-level strategy decisions. For instance, when $r=1$, the value obtained by actor i is the same as obtained directly from the design exploration task, i.e. $f_i^{\bar{s}}$; under these circumstances, agreeing on a collective design $d_{\text{sym}}^{\bar{s}}$ can be challenging if there are other local value-maximizing designs elsewhere in the design space. On the other end, setting $r=0.5$ provides each actor with the arithmetic mean of individual values. Since the actors are exposed to the same value conditions,

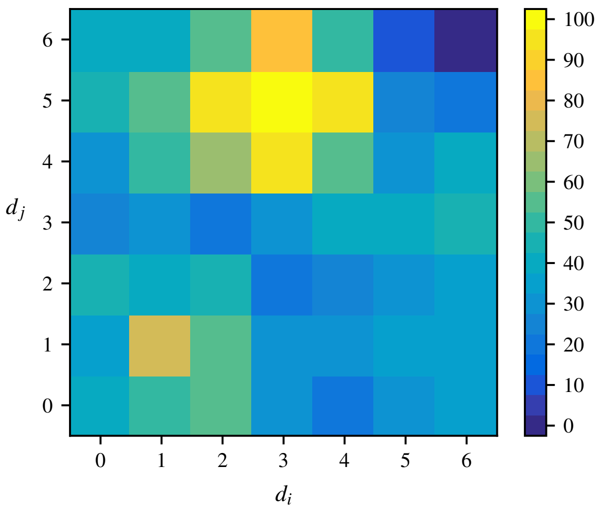


FIGURE 1. EXAMPLE VALUE SPACE $\mathbf{F}_{ij}^{\mathcal{R}}$ WITH SYMMETRIC VALUE-MAXIMIZING DESIGN $d_{\text{sym}}^{\mathcal{R}} = \langle d_i^{\mathcal{R}}, d_j^{\mathcal{R}} \rangle = \langle 1, 1 \rangle$.

this condition makes upper-level design strategy tradeoffs symmetric and more static.

2.3 Strategy Model Implementation

The strategy model implementation generates the design strategy tradeoffs between the two previously-explored design spaces to which the actors are exposed. At this level, designs are treated as strategies in a normal-form game with incomplete information. If actors agree on a $\bar{s} = \langle s_i, s_j \rangle \,|\, s_i = s_j = \mathcal{S} \in \{\mathcal{R}, \mathcal{B}\}$, payoff value vector $V^s = \langle V_i^s, V_j^s \rangle$ is equal to one of the previously-evaluated value functions, $V^{\mathcal{R}} = \langle V_i^{\mathcal{R}}, V_j^{\mathcal{R}} \rangle$ and $V^{\mathcal{B}} = \langle V_i^{\mathcal{B}}, V_j^{\mathcal{B}} \rangle$. Otherwise, the elements of V^s are computed via Eqn. (10) to ensure specified strategy dynamics,

$$V_i^s = \begin{cases} V_i^{\bar{s}}, & \text{if } s_i = s_j \in \{\mathcal{R}, \mathcal{B}\} \\ S \cdot V_i^{\mathcal{R}} + (1-S) \cdot V_i^{\mathcal{B}}, & \text{if } s_i = \mathcal{R} \wedge s_j = \mathcal{B} \\ T \cdot V_i^{\mathcal{R}} + (1-T) \cdot V_i^{\mathcal{B}}, & \text{if } s_i = \mathcal{B} \wedge s_j = \mathcal{R} \end{cases} . \quad (10)$$

The values from Eqn. (10) are constrained to be between 0 and 100. The S and T parameters are the "sucker's" and "temptation-to-defect" payoffs corresponding to the payoff values obtained by the two players, one that cooperates (chooses design strategy $\mathcal{R}$) while the other defects ($\mathcal{B}$), respectively, in the classical Prisoner's Dilemma game [14]. Conceptually, S motivates decisions based on fear while T motivates decisions based on greed. Specific to this work, S and T are the off-diagonal normal-form payoffs in which agreeing on $\mathcal{R}$ results in a payoff of 1 and agreeing on $\mathcal{B}$ results in a payoff of 0 as in Table 2.

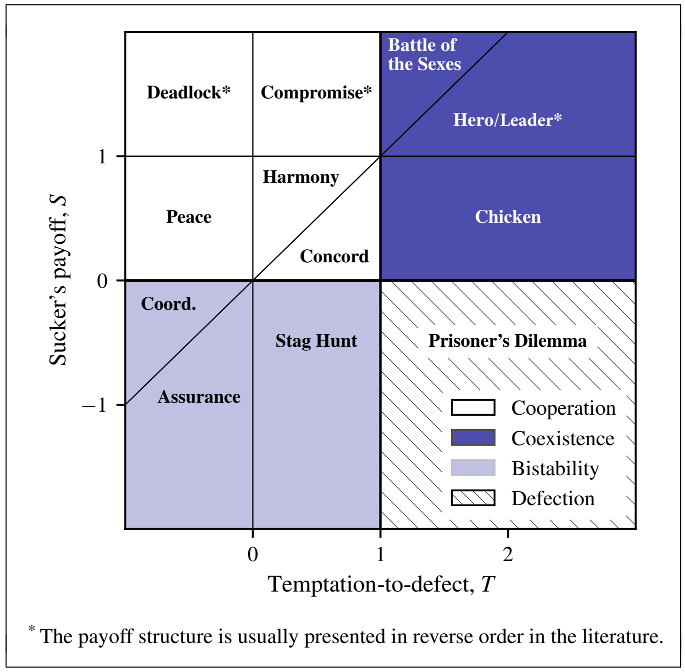


FIGURE 2. STRATEGY DYNAMICAL DOMAINS AND ASSOCIATED CLASSICAL GAMES ON THE SUCKER'S/TEMPTATION-TO-DEFECT (S–T) PLANE [14, 20]

The S–T plane can be divided in four strategy dynamical domains: cooperation, coexistence, bistability, and defection. Each domain encloses several classes of games, as illustrated in Fig. 2 [14]. Previous works have focused in the region defined by $-1 < S < 1$ and $0 < T < 2$ which contains the Harmony/Concord, Chicken (a.k.a. Snowdrift), Stag Hunt, and Prisoner's Dilemma scenarios [18, 19]. Each of these games presents a different and unique strategy dynamic: one Pareto efficient Nash equilibrium (Harmony/Concord), one non-Pareto efficient Nash equilibrium (Prisoner's Dilemma), two Nash equilibria under matching strategy decisions (Stag Hunt), and two Nash equilibria under opposite strategy decisions (Chicken).

While S and T provide some control over strategy dynamics, the resulting game depends in part on outcomes from lower-level design decisions. Logically stated, Eqn. (11) must hold for the resulting payoffs to be comparable to a classical game.

$$\left(V_i^{\mathcal{R}} > V_i^{\mathcal{B}} \wedge V_j^{\mathcal{R}} > V_j^{\mathcal{B}}\right) \oplus \left(V_i^{\mathcal{B}} > V_i^{\mathcal{R}} \wedge V_j^{\mathcal{B}} > V_j^{\mathcal{R}}\right) \quad (11)$$

Selecting designs d_{sym}^s for both the **Red** and **Blue** scenarios is sufficient to enforce Eqn. (11) and enable classification in Fig. 2. Otherwise, the resulting game may not be symmetric or may present reversed payoff values, especially if the values of S and T fall in the coexistence or the bistability domains.

TABLE 2. NORMALIZED COOPERATION/DEFECTION GAME TO CHARACTERIZE STRATEGY DYNAMICS [14], $s \in \{\mathcal{R}, \mathcal{B}\}^2$

		Player 2: Cooperate ($\mathcal{R}$)	Player 2: Defect ($\mathcal{B}$)
Player 1	Cooperate ($\mathcal{R}$)	$V_1^s = 1$, $V_2^s = 1$	$V_1^s = S$, $V_2^s = T$
	Defect ($\mathcal{B}$)	$V_1^s = T$, $V_2^s = S$	$V_1^s = 0$, $V_2^s = 0$

2.4 Limitations of the Model

The proposed CoDe model aimed to generate design decision-making problems with specified strategy dynamics carries several limitations which must be discussed. First, it treats design as a bi-level decision-making problem, with lower-level design processes represented as discipline-independent value space exploration tasks and upper-level tradeoffs synthesized from an abstract game-theoretical representation of cooperative and defective strategies. This distinction complicates design actors from sequentially or simultaneously finding value-maximizing designs and strategies as a more general design problem. However, differentiating design and strategy decisions helps to separate two timescales of decision activity: design decisions happen in a comparatively shorter timescale while strategy decisions happen on a longer timescale with greater restrictions on information sharing. Furthermore, addition of strategy control parameters (S and T) help to model desired strategy dynamics between design actors.

Second, design and strategy spaces selected in the model implementation are relatively small and characterized by abstract test functions for two actors. While not entirely representative of real-world design problems, the low-dimensional spaces help to contribute some cognitive difficulty without overwhelming design actors. Limiting cognitive difficulty is particularly important in abstract design problems which do not provide context amenable to existing knowledge and for research questions focusing on strategy dynamics.

Finally, the proposed implementation to generate the value spaces enforces symmetry between actors. The collective value-maximizing designs will always appear on the plane of symmetry, greatly simplifying the design exploration task for cognizant design actors. While most real-world design problems are not symmetric, symmetry in this design space provides a measure of control to minimize differences between actors. This limitation can be mitigated in software implementations by re-ordering the design variables to avoid the immediate appearance of symmetry.

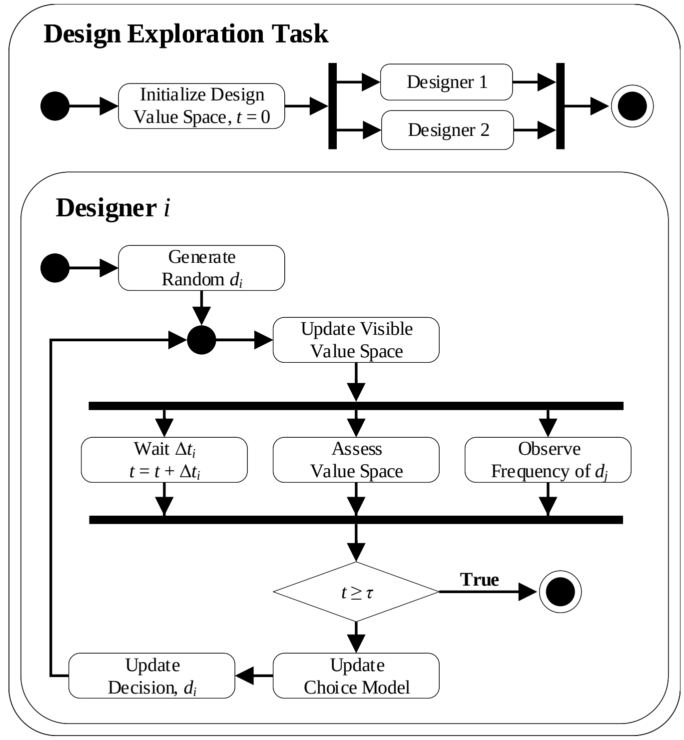


FIGURE 3. ACTIVITY DIAGRAM FOR DESIGN MODEL

3 SIMULATION STUDY

This section describes a preliminary analysis of the CoDe model described in Section 2 using multi-agent simulation and game theory. This study draws inspiration from previous works that have employed computational agents to assist CoDe research [21–23]. The next sections describe the elements of the simulation model and results from the study.

3.1 Simulation of the Design Model

The lower-level design value exploration tasks using the model implementation in Section 2.2 are studied using multi-agent simulation (Fig. 3). The design value space is initialized and designer agents 1 and 2 start to explore the design space at $d = \langle d_1, d_2 \rangle$. From this point forward, agent i's performs a set of actions every Δt_i interval of time while $t < \tau$ as follows:

1. *Update Visible Value Space*. After cell $d = \langle d_i, d_j \rangle$ is selected, the entries $f_i^{\bar{s}}(d_i, d_j)$ and $f_j^{\bar{s}}(d_j, d_i)$ are revealed to agents i and j, respectively. The technical value space visible to agent i is $\tilde{\mathbf{F}}_{ij}^{\bar{s}} = \left[\tilde{f}_i^{\bar{s}}(d_i, d_j)\right]_{|\mathbb{d}|}$, where

$$\tilde{f}_i^{\bar{s}} = \begin{cases} f_i^{\bar{s}}, & \text{if } d = \langle d_i, d_j \rangle \text{ has been observed} \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

2. *Assess Value Space*. Designer i observes $\tilde{\mathbf{F}}_{ij}^{\bar{s}}$ and estimates the hidden values to the left and right of d. If only d is visible, agent i assumes $\tilde{f}_i^{\bar{s}}(d_i+1,d_j) = \tilde{f}_i^{\bar{s}}(d_i-1,d_j) = f_i^{\bar{s}}(d_i,d_j)\ \forall\ (d_i+1),(d_i-1) \in \mathbb{d}_i$. Otherwise, the algorithm uses spline extrapolation to estimate $\tilde{f}_i^{\bar{s}} \in [0,100]$ along row d_j. The $\tilde{f}_i^{\bar{s}}$ estimates are temporarily replaced in $\tilde{\mathbf{F}}_{ij}^{\bar{s}}$.
3. *Observe Frequency of d_j*. Via the visible value space, Designer i observes j's frequency of decisions d_j as vector $\mathbf{p}_j$.
4. *Update Choice Model*. Agent i's representative utility from making $q \in \{0,\dots,K\}$ his next d_i is modeled as:

$$\mathbf{u}_i\left(t,\tilde{\mathbf{F}}_{ij}^{\bar{s}}\right) = \left[U_{d_i \leftarrow q}\left(t,\tilde{\mathbf{F}}_{ij}^{\bar{s}}\right)\right] = \left(\mathbf{p}_j\tilde{\mathbf{F}}_{ij}^{\bar{s}}\right)^{\circ\frac{\tau}{\tau-t}}, \qquad (13)$$

where the product $\mathbf{p}_j\tilde{\mathbf{F}}_{ij}^{\bar{s}}$ is a vector of the anticipated contribution by agent i from agreeing with j on $\langle d_i \leftarrow q, d_j\rangle$. The parameter $\tau/(\tau-t)$ penalizes each element of $\mathbf{p}_j\tilde{\mathbf{F}}_{ij}^{\bar{s}}$ (via Hadamard exponentiation, $\circ$) increasing the relative differences between $U_{d_i\leftarrow q}\left(t,\tilde{\mathbf{F}}_{ij}^{\bar{s}}\right)$ estimates as time advances. It also attempts to mimic the effects of time constraints on collective action, and it is based on empirical assumptions inspired by previous psychology research on the cognitive mechanics of collective decision-making processes, specifically, reported evidence of the positive effect of time pressure on cooperation [24, 25]. Finally, the choice probability mass function for choosing $d_i = q$ is computed as

$$\mathbf{p}_i = \left[\Pr\left(d_i \leftarrow q\,|\,t,\tilde{\mathbf{F}}_{ij}^{\bar{s}}\right)\right] = \frac{\mathbf{u}_i\left(t,\tilde{\mathbf{F}}_{ij}^{\bar{s}}\right)}{\left|\mathbf{u}_i\left(t,\tilde{\mathbf{F}}_{ij}^{\bar{s}}\right)\right|} \qquad (14)$$

5. *Update Decision*. The next value of d_i is selected by using inverse transform sampling on the cumulative distribution function obtained from $\mathbf{p}_i$.

Table 3 shows sequential example of how the multi-agent simulation performs the design value exploration task using constant $\Delta t_0 = \Delta t_1 = 1$ and $\tau = 100$. The initial $d = \langle d_1, d_2\rangle = \langle 0,5\rangle$ appears attractive for Designer 1 and unfavorable for Designer 2. As reflected by the number of visible cells at each d_i, Designer 1 remains more attracted to $d_1 \in \{0,1\}$ by $t = 22$ while Designer 2 is more attracted to $d_2 \in \{2,4\}$. By the end of the task, Designer 2 focuses on $d_2 = 2$, leading Designer 1 to select $d_1 = 2$, and converging to $d = d_{\text{sym}}^{\bar{s}}$.

Table 4.b shows results of a batch multi-agent simulation execution with 10,000 seeds to explore design value spaces for the **Red** and **Blue** scenarios. In both cases, $d = d_{\text{sym}}^{\bar{s}}$ was the most frequent outcome, 27.46% for the **Red** design and 20.04% for the **Blue** design. In addition, Table 4.c shows the distribution of $f^{\bar{s}}(d) = \langle f_i^{\bar{s}}, f_j^{\bar{s}}\rangle$. Consistent with the distribution of design decisions, the most frequent value outcome from the **Red** and **Blue** scenarios are $f_{\text{sym}}^{\mathcal{R}} = 75$ and $f_{\text{sym}}^{\mathcal{R}} = 55$, respectively.

3.2 Strategy Analysis

This section analyzes the implementation of the upper-level design strategy selection problem described in Section 2.3 using the results obtained from the multi-agent simulation in Section 3.1. Equation (10) computes the design strategy game in Table 1 from the obtained values $\langle V_i^{\mathcal{R}}, V_j^{\mathcal{R}}\rangle$ and $\langle V_i^{\mathcal{B}}, V_j^{\mathcal{B}}\rangle$. This section explores the Harmony/Concord and Prisoner's Dilemma strategy games respectively from the cooperation and defection dynamical domains defined by $S \in \{-1/2, 1/2\}$ and $T \in \{1/2, 3/2\}$ in Fig. 2.

Nash equilibria (N.E.) represent rational outcomes from the design strategy selection phase. Table 5 shows the distribution of single N.E. payoffs for both strategy game scenarios and the count of total resulting games with one or multiple N.E. for every possible Harmony/Concord or Prisoner's Dilemma game (i.e. $|\mathbb{d}|^2 = 2,401$) using reward mechanisms with $r \in \{1, 0.75, 0.5\}$, computed using the Nashpy Python package.

4 DISCUSSION

4.1 Insights from the Simulation Study

Results from the simulation study expose some of the challenges regarding the topology and exploration of generated value spaces. First, as the difference in values between the symmetric goal at $d_{\text{sym}}^{\bar{s}}$ and the global maximum at $d_i^\star$ grows, so do the range of obtained values $f^{\bar{s}}(d) = \langle f_i^{\bar{s}}, f_j^{\bar{s}}\rangle$. A value-maximizing agent will be attracted to visible global maximizers $d^\star$, but some of these maxima are undesirable for other agents. For instance, the global maximum $d^\star = \langle 6,0\rangle$ in design value space $\mathcal{B}$ yields $f_i^{\bar{s}} = 100$ to agent i and $f_j^{\bar{s}} = 0$ for agent j. This outcome can only be collectively efficient if a reward mechanism is used, otherwise the agents will disagree on $d^\star$ due to extreme value asymmetry. In design value space $\mathcal{R}$, the global maximum $d^\star = \langle 5,3\rangle$ yields $f_i^{\bar{s}} = 100$ to agent i and $f_j^{\bar{s}} = 40$ for agent j. In this case, however, the agents converge more frequently to decisions in close proximity to $d^\star$, in particular $d^\star = \langle 6,3\rangle$, at which agent i obtains $f_i^{\bar{s}} = 85$ and j gets $f_j^{\bar{s}} = 45$, respectively, a difference of -15 and $+5$ with respect to one agent reaching the global maximizer.

In the game-theoretical analysis, a higher r results in a more scattered distribution of N.E. away from the diagonal payoffs $V_i^s = V_j^s$. When $r > 0.5$, most of the resulting games are asymmetric with N.E. located on the off-diagonal of the design strategy game in Table 1. In the Harmony/Concord scenario, the most common N.E. payoff is $V_i^s = V_j^s = 75 = V_{\text{sym}}^{\mathcal{R}}$, whose frequency remains the same for different values of r and it is also a Pareto efficient solution. On the other hand, for the Prisoner's Dilemma games, the most frequent N.E. payoff goes from $V_i^s = V_j^s \simeq 40$

TABLE 3. SIMULATED VALUE SPACE EXPLORATION FOR STRATEGY $\mathcal{B}$ WITH CONSTANT $\Delta t = 1$ AND $\tau = 100$, CONVERGING TO $d^{\bar{s}}_{\text{sym}}$ AT $t = 86$.

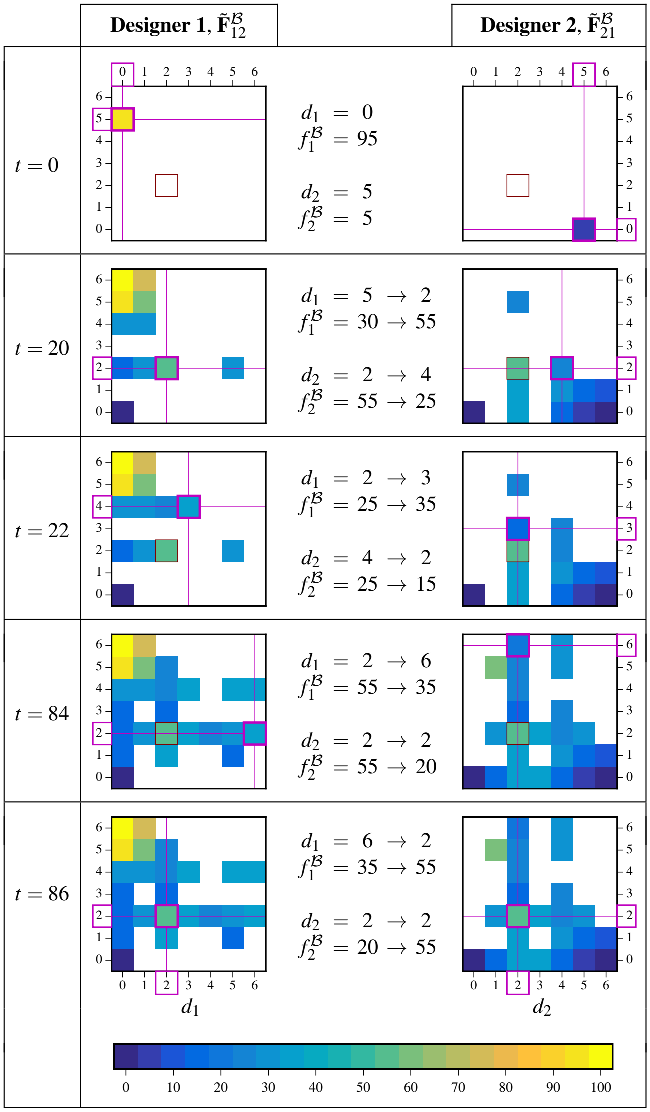
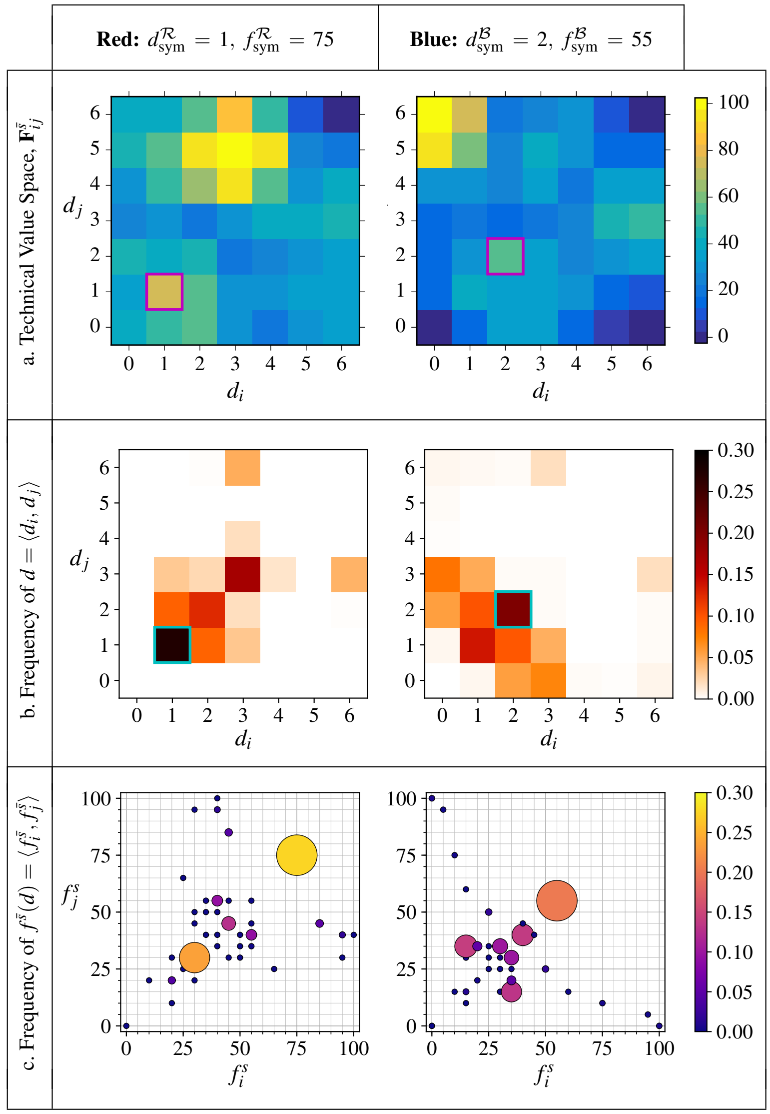


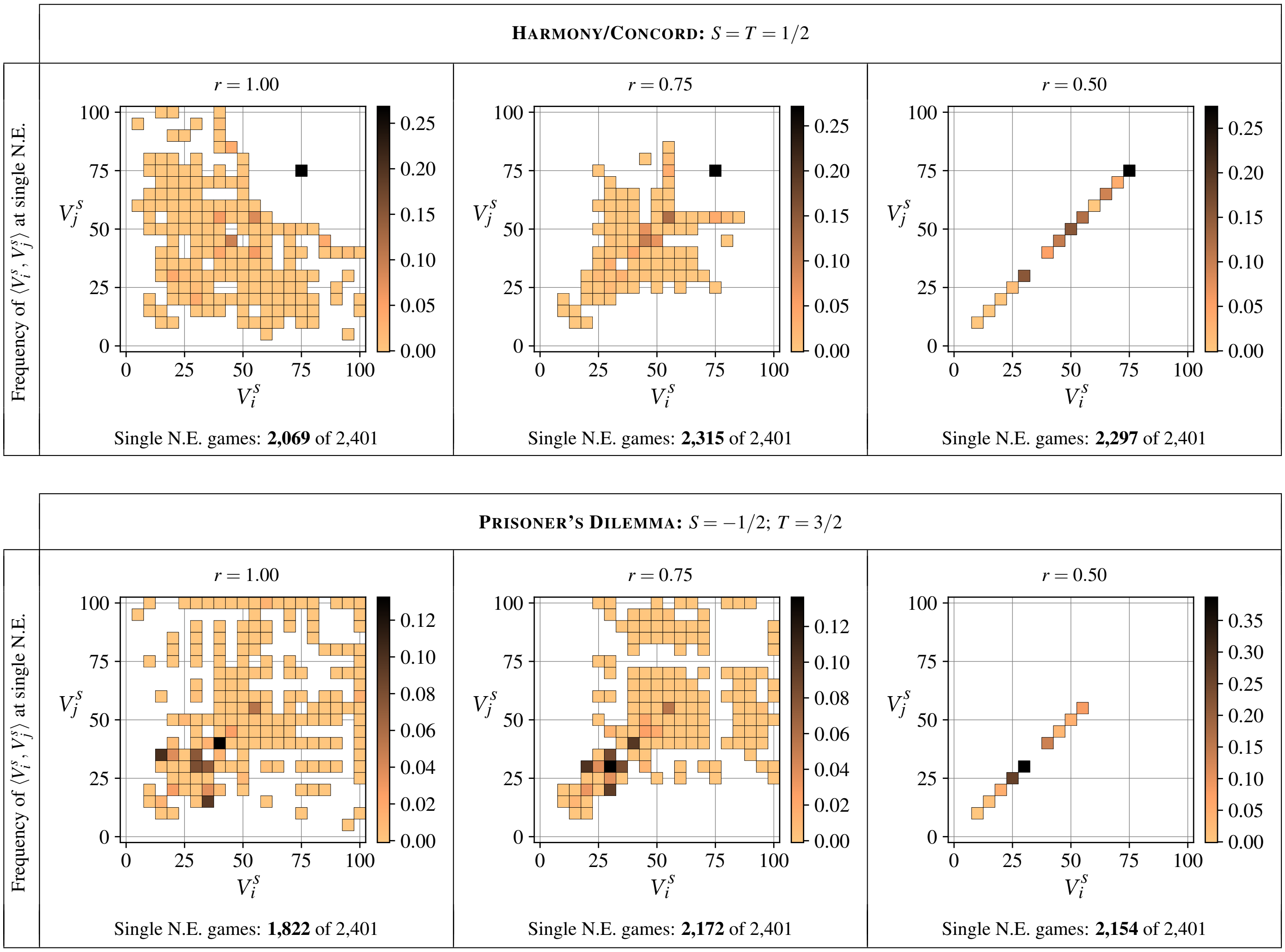
TABLE 4. RESULTS FROM A MULTI-AGENT SIMULATION BATCH EXECUTION OF THE DESIGN MODEL USING 10,000 DIFFERENT SEEDS FOR TWO DESIGNS, $\mathcal{R}$ AND $\mathcal{B}$



for $r = 1.00$ to $V^s_i = V^s_j \simeq 30$ for $0.50 < r < 1.00$. While in the Harmony/Concord scenario the highest single N.E. payoff at which both players can agree is $V^{\mathcal{R}}_{\text{sym}}$, the Prisoner's Dilemma design tradeoff yields single N.E. rewards of up to $V^s_i = V^s_j = 100$. However, these outcomes are not reached by the artificial agents described in Fig. 3, and the possibility of reaching them disappears for lower values r. In general, a reward mechanism with $r = 0.50$ eliminates any chances of having improved collective outcomes under the strategy dynamics considered.

These observations generate new research questions to be addressed in future experimentation with human subjects. Particularly, 1) what characteristics of the design value spaces affect outcomes of multi-agent design exploration tasks, 2) how do collective reward mechanisms affect design actors' exploration of the value space, 3) are actors that agree on fairer decisions in the exploration tasks more likely to cooperate in a design game under specific strategy dynamics, 4) how do collective reward mechanisms in the design process affect collective effort in strategy selection, and 5) under which conditions can actors reach payoff values that surpass results obtained by artificial design agents?

TABLE 5. PAYOFF VALUES $\langle V_i^s, V_j^s \rangle$ AT SINGLE NASH EQUILIBRIA (N.E.) FOR $s_i, s_j \in \{\mathcal{R}, \mathcal{B}\}$ UNDER TWO DIFFERENT STRATEGY GAME SCENARIOS, HARMONY/CONCORD AND PRISONER'S DILEMMA



4.2 Limitations of the Study

The simulation study is a first step to explore and study the usefulness and extensiveness of the proposed CoDe model and is subject to several limitations which must be acknowledged. First, the designer behavior algorithm is unidirectional, iterating through one technical design decision space at a time. The algorithm does not iterate between the different value spaces or between lower-level technical and upper-level design strategy decisions. This limitation makes it difficult for agents to make collective design and strategy decisions that result in maximum payoff values.

Second, the algorithm is a simplified model of decision-making. It assumes limited knowledge of the design domain and does not allow communication between players beyond visibility of cells in the value space. It accounts for utility-maximizing decision-making agents with representative utility expressed in terms of the frequency of both agents' decisions. Nevertheless, as the algorithm relies on stochastic behavior to model variability in design decisions, the utility model is subject to bias towards individual value-maximizing rewards rather than symmetric goals. A more realistic decision-algorithm may take into account interactive effects between agents, for example by including reward mechanisms to the simulation, building on existing theory of iterative negotiation [26, 27].

Finally, the game-theoretical analysis of the design tradeoff model only focuses on two particular strategy game scenarios to highlight divergent outcomes under fixed decision algorithms. These scenarios were selected from a broader investigation of the strategy dynamical domains in Fig. 2. This work is limited to assess design strategy tradeoffs with only one N.E., characteristic of games in cooperation and defection domains. Games resulting from the coexistence and bistability scenarios are mostly bipolar (two pure N.E.), and each strategy must be analyzed in terms of payoff and risk dominance [13, 28], which is out of the scope for this work. Future analyses will work to consider a broader range of mechanisms and strategy dynamics to gain more insights on the challenges of collective systems design.

4.3 Implications for Collective Systems Design

This work provides insights on how the structure of a multi-actor design problem can evoke different strategy dynamics between independent decision-makers. The value spaces are not inherently fixed and can be shaped to align actors with desirable outcomes. For instance, the constraint assumptions in Eqn. (5) to (8) can be modified to make the global maxima coincide with the symmetric goals or assign a higher value to the latter to make it more explicit. Also, improved mechanisms and incentives may overcome misaligned objectives between actors. Future work seeks to further classify multi-actor design problems into a set of prototypical designs recognizing the inherent dynamics among actors. This work will contribute to design of complex systems by identifying and avoiding poorly-structured design problems.

5 CONCLUSIONS AND FUTURE WORK

This work contributes to literature in engineering systems design by proposing a model-based approach to study design decisions and strategic behavior among independent actors. The proposed CoDe framework builds upon a multi-actor value model and game theory to explore the dynamics of collective design in a bi-level decision-making process. Specifically, the CoDe activity differentiates between two phases: exploration of discipline-independent design value spaces and strategy tradeoffs. The first phase addresses how technical aspects of a design problem (e.g. size, coupling between decisions) influence collective action. The second phase frames the decision-making process at the organizational level as a two-strategy normal-form game with tradeoffs synthesized from an abstract game-theoretical representation of cooperative and defective strategies. This approach helps generate synthetic design problems with specified strategy dynamics that can be used in future experimentation with cognizant design actors.

Future work will conduct a human designer experiment based on the CoDe model to assess the technical and social factors impacting actual collective decision-making processes in engineering systems design subject to biases, heuristics, and satisficing behaviors. The longer-term goal of this research is to help understand what design problem structures are suitable for collective efforts and how to effectively organize actors to handle distributed and federated design. Going in this direction, this research aims to impact the way large engineering projects are viewed to consider multiple collective perspectives and the potential role of coordination mechanisms to achieve desired behavior.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. CMMI–1742971 titled “EAGER: Model-based Foundations of Collective Systems Design Theory.”

REFERENCES

[1] Gurnani, A., and Lewis, K., 2008. “Collaborative, Decentralized Engineering Design at the Edge of Rationality”. *Journal of Mechanical Design,* **130**(12), p. 121101.

[2] Collopy, P. D., and Hollingsworth, P. M., 2011. “Value-driven Design”. *Journal of Aircraft,* **48**(3), pp. 749–759.

[3] Martins, J. R. R. A., and Lambe, A. B., 2013. “Multidisciplinary Design Optimization: A Survey of Architectures”. *AIAA Journal,* **51**(9), pp. 2049–2075.

[4] Cao, Y., Yu, W., Ren, W., and Chen, G., 2013. “An Overview of Recent Progress in the Study of Distributed Multi-Agent Coordination”. *IEEE Transactions on Industrial Informatics,* **9**(1), pp. 427–438.

[5] Lu, S.-Y., ElMaraghy, W., Schuh, G., and Wilhelm, R., 2007. “A Scientific Foundation of Collaborative Engineering”. *CIRP Annals-Manufacturing Technology,* **56**(2), pp. 605–634.

[6] Iorio, J., and Taylor, J. E., 2014. “Boundary Object Efficacy: The Mediating Role of Boundary Objects on Task Conflict in Global Virtual Project Networks”. *International Journal of Project Management,* **32**(1), p. 121101.

[7] Klein, M., Faratin, P., Sayama, H., and Bar-Yam, Y., 2003. “Negotiating Complex Contracts”. *Group Decision and Negotiation,* **12**(2), pp. 111–125.

[8] de Kraker, J., Croeze, C., and Kirschner, P., 2011. “Computer Models as Social Learning Tools in Participatory Integrated Assessment”. *International Journal of Agriculture Sustainability,* **9**(2), pp. 297–309.

[9] Reich, Y., 2010. “My Method is Better!”. *Research in Engineering Design,* **21**(3), pp. 137–142.

[10] Sha, Z., Kannan, K. N., and Panchal, J. H., 2015. “Behavioral Experimentation and Game Theory in Engineering Systems Design”. *Journal of Mechanical Design,* **137**(5), p. 051405.

[11] Bhatia, G. V., Kannan, H., and Bloebaum, C. L., 2016. “A Game Theory Approach to Bargaining over Attributes of

Complex Systems in the Context of Value-Driven Design". In *Proceedings of the 54th AIAA Aerospace Sciences Meeting*, San Diego, CA, USA, p. 0972.
[12] Papageorgiou, E., Eres, M. H., and Scanlan, J., 2016. "Value Modelling for Multi-stakeholder and Multi-objective Optimisation in Engineering Design". *Journal of Engineering Design,* **27**(10), pp. 697–724.
[13] Grogan, P. T., Ho, K., Golkar, A., and de Weck, O. L., 2018. "Multi-actor Value Modeling for Federated Systems". *IEEE Systems Journal,* **12**(2), pp. 1193–1202.
[14] Hauert, C., 2002. "Effects of Space in 2×2 Games". *International Journal of Bifurcation and Chaos,* **12**(07), pp. 1531–1548.
[15] Miller, S. W., Simpson, T. W., Yukish, M. A., Stump, G., Mesmer, B. L., Tibor, E. B., Bloebaum, C. L., and Winer, E. H., 2014. "Toward a Value-driven Design Approach for Complex Engineered Systems using Trade Space Exploration Tools". In *Proceedings of the ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Buffalo, NY, USA, pp. V02AT03A052–V02AT03A052.
[16] Isaksson, O., Bertoni, M., Hallstedt, S., and Lavesson, N., 2015. "Model Based Decision Support for Value and Sustainability in Product Development". In *Proceedings of the 20th International Conference on Engineering Design*, Milan, Italy, Vol. 1, The Design Society, pp. 21–30.
[17] Yang, X.-S., 2010. "Firefly Algorithm, Stochastic Test Functions and Design Optimisation". *International Journal of Bio-Inspired Computation,* **2**(2), pp. 78–84.
[18] Roca, C. P., Cuesta, J. A., and Sánchez, A., 2009. "Evolutionary Game Theory: Temporal and Spatial Effects Beyond Replicator Dynamics". *Physics of Life Reviews,* **6**(4), pp. 208–249.
[19] Gianetto, D. A., and Heydari, B., 2016. "Sparse Cliques Trump Scale-free Networks in Coordination and Competition". *Scientific Reports,* **6**.
[20] Bruns, B. R., 2015. "Names for Games: Locating 2×2 Games". *Games,* **6**(4), pp. 495–520.
[21] Vermillion, S. D., and Malak, R. J., 2018. "A Game Theoretical Perspective on Incentivizing Collaboration in System Design". In *Disciplinary Convergence in Systems Engineering Research*, A. M. Madni, B. Boehm, R. G. Ghanem, D. Erwin, and M. J. Wheaton, eds. Springer, ch. 59, pp. 845–855.
[22] McComb, C., Cagan, J., and Kotovsky, K., 2015. "Lifting the Veil: Drawing Insights about Design Teams from a Cognitively-inspired Computational Model". *Design Studies,* **40**, pp. 119–142.
[23] Crowder, R. M., Robinson, M. A., Hughes, H. P., and Sim, Y.-W., 2012. "The Development of an Agent-based Modeling Framework for Simulating Engineering Team Work". *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans,* **42**(6), pp. 1425–1439.
[24] Cone, J., and Rand, D. G., 2014. "Time Pressure Increases Cooperation in Competitively Framed Social Dilemmas". *PLOS One,* **9**(12), p. e115756.
[25] Rand, D. G., Greene, J. D., and Nowak, M. A., 2012. "Spontaneous Giving and Calculated Greed". *Nature,* **489**, pp. 427–430.
[26] Klein, M., Sayama, H., Faratin, P., and Bar-Yam, Y., 2006. "The Dynamics of Collaborative Design: Insights from Complex Systems and Negotiation Research". In *Complex Engineered Systems*, D. Braha, A. A. Minai, and Y. Bar-Yam, eds. Springer, ch. 8, pp. 158–174.
[27] Sayama, H., Farrell, D. L., and Dionne, S. D., 2011. "The Effects of Mental Model Formation on Group Decision Making: An Agent-based Simulation". *Complexity,* **16**(3), pp. 49–57.
[28] Selten, R., 1995. "An Axiomatic Theory of a Risk Dominance Measure for Bipolar Games with Linear Incentives". *Games and Economic Behavior,* **8**(1), pp. 213–263.