# Cooperative Max Games and Agent Failures<sup>\*</sup>

Yoram Bachrach Microsoft Research Cambridge, UK vorambac@gmail.com Rahul Savani University of Liverpool, UK rahul.savani@liverpool.ac.uk

Nisarg Shah Carnegie Mellon University, USA nkshah@cs.cmu.edu

# ABSTRACT

Procuring multiple agents with different ability levels to independently solve the same task is common in labor markets, crowdsourcing environments and research and development projects due to two reasons: some agents may fail to provide a satisfactory solution, and the redundancy increases the quality of the best solution found. However, incentivizing large number of agents to compete for one task is difficult; agents need fair ex-ante guaranteed payoffs that consider their ability levels and failure rates to exert efforts.

We model such domains as a cooperative game called the Max-Game, where each agent has a weight representing its ability level, and the value of an agent coalition is the maximal weight of the agents in the coalition. When agents may fail, we redefine the value of a coalition as the expected maximal weight of its surviving members. We analyze the core, the Shapley value, and the Banzhaf index as methods of payoff division. Surprisingly, the latter two, which are usually computationally hard, can be computed in polynomial time. Finally, we initiate the study of a new form of sabotage where agents may be incentivized to influence the failure probabilities of their peers, and show that no such incentive is present in a restricted case of Max-Games.

### **Categories and Subject Descriptors**

I.2.11 [Distributed Artificial Intelligence]: Multiagent systems; J.4 [Social and Behavioral Sciences]: Economics

#### **General Terms**

Algorithms, Economics

### Keywords

Cooperative game theory, Agent failures, Max games

# 1. INTRODUCTION

Consider developing a bank of blood or organ donors. Different donations may be of different quality. For example,

\*The second author is supported by ESRC grant ESRC/BSB/09 and EPSRC grant EP/L011018/1.

Appears in: Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France. Copyright © 2014, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved. a donation from an old donor may be usable for less time than a donation from a young donor, and a donation from a completely healthy person is more desirable than a donation from a person with minor illnesses. Further, having a person in the bank does not guarantee with absolute confidence that a donation could be made successfully when the need arises. The patient and the potential donor may have blood type or tissue type incompatibilities, preventing a donation. Some people, e.g., those with blood type "O negative", have a higher probability of being compatible donors than others.

Potential donors can thus be characterized by both their qualities and their probabilities of being able to successfully donate when required. In that case, a large centralized donor bank is very desirable since it increases both the probability of finding a match, and the quality of the best match found. However, many people are reluctant to participate due to the hassle involved. It is essential to put efforts in incentivizing the right people that are the most beneficial to the bank. Who are the most important donors? How do we incentivize them fairly?

Alternatively, consider a firm researching multiple alternative solutions for a task, e.g., multiple technologies for new product development. The firm may hire several teams, each independently researching a different technology. Research is risky, so a team researching a technology may fail to develop it. In the end, the firm would choose the best technology among the ones successfully developed, which determines the utility to the firm. However, it is not wise to reward only the team that developed the chosen technology. Such a reward scheme would deter teams that are researching riskier but highly suited technologies from putting in effort in the first place, as their efforts may ultimately be in vain. One approach is to pre-determine the reward to each team based on the expected contribution of its efforts. Can we measure such contributions and fairly determine the exante rewards? Can the teams themselves come to a stable ex-ante agreement for dividing the final joint payoff?

Viewing each research team as an agent, the crux of the above hypothetical scenario is that the utility to the principal is determined by the maximum quality of solutions delivered by agents that did not fail. This is commonplace in the real world. For example, in crowdsourcing marketplaces such as Amazon's Mechanical Turk or oDesk, requesters sometimes hire multiple workers for the same job, since several workers may fail to deliver a satisfactory solution, and also to increase the maximum quality among the received solutions [16, 14]. It is common to guarantee the workers a certain payment in advance to incentivize them to exert sincere efforts. Once again, fair ex-ante payments should consider the skill levels of the workers and their success rate in delivering satisfactory solutions, both of which may be estimated based on past performance.

Similarly, in *all-pay auctions* and *crowdsourcing contests*, several teams are invited to compete to solve a task. For instance, the media provider Netflix issued a \$1,000,000 prize in a contest to improve its movie recommender system [11]. Many teams submitted their recommendation algorithms, and Netflix chose the winner to replace its old algorithm. A similar business model is employed by Topcoder and CodeChef, who organize programming contests: participants submit code for a specified problem, and compete for rewards. While such contests do not offer purely ex-ante payoffs, they vaguely resemble Max-Games in that they offer prizes to more than one team to incentivize a large number of teams to participate, in turn increasing the quality of the best solution developed.

In the above examples, redundancy compensates for uncertain agent failures, but highlights the difficulty of incentivizing agents to exert effort. Increasing the number of participating agents increases the overall success probability of the project and the expected utility. On the other hand, self-interested agents would only exert effort when allocated a high enough share of the resulting profits. In view of this inevitable tradeoff, it is essential to find a fair, and hopefully stable, reward-sharing scheme that takes into account both agents' failure probabilities and the quality levels they would provide if they succeeded.

Further, suppose that agents agree on a reward sharing scheme they believe fairly reflects individual contributions. In various domains, agents possess partial or full power to sabotage their peers by increasing their failure probabilities. For example, agents can hide information or limit their support; in a network setting agents may stop some of the traffic or fail to maintain parts of the network to increase the risk of communication errors, and in a multi-sensor network agents may withhold some of the readings from their sensors, making it harder for other agents to detect changes in the environment. Increasing failure probabilities reduces the overall expected quality of the joint project, diminishing the total reward, but also changes the agent's share of the total reward obtained. Would agents have an incentive sabotage their peers in order to increase their own reward?

We investigate the above questions by modeling the underlying key aspect of such domains as a cooperative game called the *Max-Game*. First, we propose a base Max-Game *without agent failures*, where each agent has a weight representing its quality or the utility it can achieve, and the value of a coalition is the maximal weight of the agents in it. To model agent failures, we use the reliability extension model [4], where each agent has a probability of "surviving", agent survivals are independent, and the value of a coalition is the expected value of its surviving sub-coalition.

**Our Contribution:** We provide results for general reliability extensions of Max-Games, as they subsume the special case of no agent failures. We show that reliability extensions of Max-Games are monotonic and submodular. On the negative side, these games have an empty core, meaning that stable payoff divisions are not possible. We demonstrate that failures actually help reduce agents' resistance to cooperation, quantifying it through the Cost of Stability [2]. On the positive side, we provide polynomial-time algorithms to

compute fair payoff divisions given by the Shapley value and the Banzhaf index. Finally, we examine incentives to agents for manipulating failure probabilities of their peers when using the Shapley value to share the rewards. We show that although in general every agent is better off increasing the failure probability of every other agent, in the restricted case of uniform failure probability no agent is incentivized to increase the common failure probability.

#### **1.1 Preliminaries**

A transferable utility cooperative game G = (N, v) is composed of a set of agents  $N = \{1, 2, ..., n\}$ , called the grand coalition, and a characteristic function  $v : 2^N \to \mathbb{R}$  mapping any coalition (agent subset) into the utility they can achieve together. By convention,  $v(\emptyset) = 0$ . For any agent  $i \in N$  and coalition  $S \subseteq N$ , we denote  $S \cup \{i\}$  by S + iand  $S \setminus \{i\}$  by S - i. A cooperative game is *monotonic* if for any  $S \subseteq T \subseteq N$ , we have  $v(S) \leq v(T)$ . A game is *submodular* if for any  $i \in N$  and any  $S \subseteq T \subseteq N - i$ , we have  $v(S+i)-v(S) \geq v(T+i)-v(T)$ . An equivalent formulation of submodularity is that  $v(S) + v(T) \geq v(S \cup T) + v(S \cap T)$ for all coalitions  $S, T \subseteq N$ . A game is *subadditive* if  $v(S) + v(T) \geq v(S \cup T)$  for all coalitions  $S, T \subseteq N$ . Submodularity implies subadditivity, since  $v(S \cap T) > 0$ .

**The Core:** The characteristic function gives the value that a coalition achieves, but not how it should be *distributed* among its members. For a payment vector  $\mathbf{p} = (p_1, \ldots, p_n)$ ,  $p_i$  denotes the payoff to agent i and  $p(C) = \sum_{i \in C} p_i$  denotes the joint payoff to a coalition C. The core requirement for a payoff vector is that the payoffs should be a distribution of the total gains, and that every coalition must receive at least as much as it can gain on its own, so no coalition can gain by defecting from the grand coalition of all agents. The core [15] is defined as the set of all imputations  $\mathbf{p}$  such that p(N) = v(N) and  $p(S) \geq v(S)$  for all  $S \subseteq N$ . The core may be empty, or may contain finitely or infinitely many imputations.

The Cost of Stability: In games where the core is empty, it is impossible to distribute the grand coalition's gains in a stable way. An external party may incentivize cooperation by offering a supplemental payment if the grand coalition is formed. Bachrach et. al. [2] formalized this as follows. Given a game G = (N, v) and a supplemental payment  $\Delta \in \mathbb{R}_+$ , the adjusted game  $G(\Delta) = (N, v')$  has characteristic function defined by:  $v'(N) = v(N) + \Delta$  and v'(S) = v(S)for  $S \neq N$ . The Cost of Stability (CoS) of a game G, denoted CoS(G), is the minimum supplemental payment  $\Delta^*$ for which the core of the adjusted game  $G(\Delta^*)$  is non-empty. The CoS quantifies the extent of instability in a game by measuring the subsidy required to overcome agents' resistance to cooperation.

**Power Indices:** While the core focuses on ways to divide the gains based on stability, power indices analyze the contributions of the agents to different coalitions, proposing ways to divide the gains based on fairness criteria. The marginal contribution of an agent i to a coalition  $S \subseteq N - i$ is v(S + i) - v(S). The Banzhaf index  $\beta_i$  of agent i is its average marginal contribution to all coalitions that do not contain it [9], that is,

$$\beta_i = \frac{1}{2^{n-1}} \sum_{S \subseteq N-i} (v(S+i) - v(S)).$$
(1)

Another power index is the Shapley value, which is uniquely

characterized by four important fairness axioms [28]. For any permutation  $\pi$  of agents, let  $\Gamma_i^{\pi} = \{j | \pi(j) < \pi(i)\}$  be the set of agents before agent i in  $\pi$ . The Shapley value of agent i, denoted  $\phi_i$ , is given by:  $\phi(i) = \frac{1}{n!} \sum_{\pi \in S_n} (v(\Gamma_i^{\pi} + i) - v(\Gamma_i^{\pi}))$ . For any coalition  $S \subseteq N - i$ , the number of permutations  $\pi \in S_n$  where  $\Gamma_i^{\pi} = S$  is exactly  $(|S|)! \cdot (n - |S| - 1)!$ , since agents in S and  $N \setminus S - i$  can appear before and after agent i in any of |S|! and (n - |S| - 1)! orders respectively. Thus, alternatively:

$$\phi_i = \frac{1}{n!} \sum_{S \subseteq N-i} \left[ |S|! (n - |S| - 1)! (v(S + i) - v(S)) \right].$$
(2)

**Reliability Games:** A model for agent failures in cooperative games was proposed in [4]. A reliability game  $G^{\mathbf{r}} = (N, v, \mathbf{r})$  consists of the set of agents  $N = \{1, 2, ..., n\}$ , the base characteristic function  $v : 2^N \to \mathbb{R}$  describing values in the absence of failures, and the reliability vector  $\mathbf{r}$ where  $r_i$  is the probability of agent *i* surviving (i.e., not failing). The characteristic function  $v^{\mathbf{r}}$  of  $G^{\mathbf{r}}$  is the expected value of the survivors: for any coalition  $S \subseteq N$ ,

$$v^{\mathbf{r}}(S) = \sum_{S' \subseteq S} \left( \prod_{i \in S'} r_i \cdot \prod_{j \in S \setminus S'} (1 - r_j) \right) \cdot v(S').$$
(3)

Here,  $\Pr[S'|S] = \prod_{i \in S'} r_i \cdot \prod_{j \in S \setminus S'} (1-r_j)$  is the probability that every agent in S' survives and every agent in  $S \setminus S'$  fails, so  $v^{\mathbf{r}}(S)$  is the expected utility S achieves under failures. For the base game G = (N, v), the game  $G^{\mathbf{r}} = (N, v, \mathbf{r})$  is called the *reliability extension* of G with reliability vector  $\mathbf{r}$ .

## 2. OUR MODEL

We first introduce a new cooperative game called the *Max-Game*, where each agent has an associated weight, and the value of a coalition is the *maximal* weight of agents in it.

DEFINITION 1 (MAX-GAME). A Max-Game is denoted by  $G = (N, \mathbf{w})$ , where  $N = \{1, 2, ..., n\}$  is the set of agents, and  $\mathbf{w} = (w_1, w_2, ..., w_n)$  is the vector of agent weights with  $w_i > 0$  for all  $i \in N$ . The characteristic function v of the game is given by  $v(S) = \max_{i \in S} w_i$ , for every  $S \subseteq N$ .

Agent weights are assumed to be positive since zero weight agents do not contribute to any coalition and can be ignored. WLOG, we also assume that agents are sorted by weight, so  $w_i \ge w_{i+1}$  for all  $i \in \{1, \ldots, n-1\}$ . Ties among equal weight agents are broken arbitrarily; our results hold irrespective of the tie-breaking used. For any coalition S, let  $min(S) = min\{i \in S\}$  denote the agent with the smallest index in S. Note that agent min(S) has the highest weight among all agents in S. Max-Game models situations discussed in Section 1, where a coalition is as strong as its strongest member. We now give a formal definition of Max-Games with agent failures using the *reliability extension model* [4] (see Section 1.1).

DEFINITION 2 (MAX-GAME WITH FAILURES). A reliability extension of a Max-Game  $G = (N, \mathbf{w})$  is denoted by  $G^{\mathbf{r}} = (N, \mathbf{w}, \mathbf{r})$ , where  $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ , and  $r_i$  is the probability of agent *i* surviving.

We discuss Max-Games without failures, but provide results directly for Max-Games with failures as they subsume results for the former with reliability vector  $\mathbf{r} = (1, ..., 1)$ .

Theorem 1 (Section 3) implies that Max-Games without failures have an empty core, meaning that cooperation does not arise naturally in such games. This is not surprising. Indeed, the maximum weight agent in any coalition has no incentive to collaborate with the rest of the agents in that coalition, since they do not contribute to the value of the coalition but share the payoff.

When failures are present, an agent that may fail might want to collaborate even with a lower weight agent since the latter can succeed and generate value when the former fails to do so. Such collaboration still comes at the cost of sharing the resulting payoff. This tradeoff, which was absent in Max-Games without failures, makes it unclear if cooperation is feasible under failures. We show (Theorem 1) that the core is still empty in all reliability extensions of Max-Games. On the positive side, Theorems 2 and 3 imply that both the Shapley value and the Banzhaf index can be computed in polynomial time. Hence, it is practical to use them for enforcing cooperation via a fair division of the payoff.

#### 2.1 Value of a Coalition

We begin by showing that the value of a coalition in any reliability extension  $G^{\mathbf{r}}$  of a Max-Game has a simple form. Let S be a coalition and  $i^* = min(S)$ . Now,  $v^{\mathbf{r}}(S)$  is the expected value of the surviving sub-coalition of S. Agent  $i^*$ survives with probability  $r_{i^*}$ , and then the value generated is  $w_{i^*}$  irrespective of the survival of others. Agent  $i^*$  fails with probability  $1 - r_{i^*}$ , and then the expected value generated is  $v^{\mathbf{r}}(S - i^*)$ . Thus,  $v^{\mathbf{r}}(S) = r_{i^*}w_{i^*} + (1 - r_{i^*})v^{\mathbf{r}}(S - i^*)$ . Expanding  $v^{\mathbf{r}}(S - i^*)$  similarly, the coefficient of  $w_i$   $(i \in S)$ in the expansion is  $r_i \cdot \prod_{i \in S, i \leq i} (1 - r_i)$ .

in the expansion is  $r_i \cdot \prod_{j \in S, j < i} (1 - r_j)$ . For any  $C \subseteq N$  and  $i \in N$ , define  $C_b^i = \{j \in C | j < i\}$ and  $C_a^i = \{j \in C | j > i\}$ . For any  $C \subseteq N$ , also define  $D(C) = \prod_{j \in C} (1 - r_j)$ , the probability that no agent in Csurvives; D(C) = 1 if  $C = \emptyset$ . Thus, we have the following.

LEMMA 1. For any  $S \subseteq N$  and  $i^* = min(S)$ ,

$$v^{\mathbf{r}}(S) = r_{i^*} w_{i^*} + (1 - r_{i^*}) v^{\mathbf{r}}(S - i^*) = \sum_{i \in S} D(S_b^i) \cdot r_i w_i.$$

COROLLARY 1. For any  $i \in N$  and  $S \subseteq N - i$ ,

$$v^{\mathbf{r}}(S+i) - v^{\mathbf{r}}(S) = r_i \cdot D(S_b^i) \cdot (w_i - v^{\mathbf{r}}(S_a^i)).$$

PROOF. We apply Lemma 1:

$$\begin{aligned} v^{\mathbf{r}}(S+i) &- v^{\mathbf{r}}(S) \\ &= \sum_{j \in S_{b}^{i}} D(S_{b}^{j}) \cdot r_{j}w_{j} + D(S_{b}^{i}) \cdot r_{i}w_{i} \\ &+ \sum_{j \in S_{a}^{i}} \left( D(S_{b}^{i}) \cdot (1-r_{i}) \cdot D((S_{a}^{i})_{b}^{j}) \right) r_{j}w_{j} \\ &- \sum_{j \in S_{a}^{i}} D(S_{b}^{j}) \cdot r_{j}w_{j} - \sum_{j \in S_{a}^{i}} \left( D(S_{b}^{i}) \cdot D((S_{a}^{i})_{b}^{j}) \right) r_{j}w_{j} \\ &= D(S_{b}^{i}) \cdot r_{i}w_{i} + D(S_{b}^{i}) \cdot (1-r_{i})v^{\mathbf{r}}(S_{a}^{i}) - D(S_{b}^{i}) \cdot v^{\mathbf{r}}(S_{a}^{i}) \\ &= r_{i} \cdot D(S_{b}^{i}) \cdot (w_{i} - v^{\mathbf{r}}(S_{a}^{i})). \end{aligned}$$

The first transition breaks the summation in Lemma 1 into agents before i, after i and i itself. The second transition uses Lemma 1 for  $v^{\mathbf{r}}(S_a^i)$ .

Next, we show some basic properties of reliability extensions of Max-Games.

LEMMA 2. Any reliability extension of a Max-Game is monotonic and submodular, and hence subadditive.

PROOF. Let  $G^{\mathbf{r}} = (N, \mathbf{w}, \mathbf{r})$  be any reliability extension of a Max-Game, and  $v^{\mathbf{r}}$  be its characteristic function. For monotonicity, we prove  $v^{\mathbf{r}}(S+i) \geq v^{\mathbf{r}}(S)$  for all  $i \in N$ and  $S \subseteq N - i$ , by induction on k = |S|. Fix any  $i \in N$ . The base case k = 0 is trivial. For any k < n, assume  $v^{\mathbf{r}}(C+i) \geq v^{\mathbf{r}}(C)$  for every  $C \subseteq N-i$  with |C| < k. Take any  $S \subseteq N - i$  with |S| = k. Let  $i^* = \min(S + i)$ . We examine two cases:

- 1.  $i^* = i$ . Using Lemma 1, we have  $v^{\mathbf{r}}(S+i) = r_i w_i + i$  $(1-r_i) \cdot v^{\mathbf{r}}(S)$ . Now,  $w_i \geq v^{\mathbf{r}}(S)$  since every agent in S has weight at most  $w_i$ , so  $v^{\mathbf{r}}(S+i) \geq v^{\mathbf{r}}(S)$ .
- 2.  $i^* \neq i$ . Using Lemma 1, we have  $v^{\mathbf{r}}(S+i) = r_{i^*}w_{i^*} + i$  $(1 - r_{i^*})v^{\mathbf{r}}(S - i^* + i)$ . Now  $|S - i^*| = k - 1$ , and by the induction hypothesis,  $v^{\mathbf{r}}(S - i^* + i) \ge v^{\mathbf{r}}(S - i^*)$ . Thus,  $v^{\mathbf{r}}(S+i) \ge r_{i^*} w_{i^*} + (1-r_{i^*}) v^{\mathbf{r}}(S-i^*) = v^{\mathbf{r}}(S),$ where the last step uses Lemma 1.

Next, we show that any base Max-Game  $G = (N, \mathbf{w})$ (without failures) is submodular. For any  $S, T \subseteq N$ ,

$$v(S \cup T) = \max\left\{\max_{i \in S} w_i, \max_{j \in T} w_j\right\} = \max(v(S), v(T))$$

Also,  $v(S \cap T) \leq v(S)$  and  $v(S \cap T) \leq v(T)$  (monotonicity) imply  $v(S \cap T) \leq \min(v(S), v(T))$ . Thus,

$$v(S \cup T) + v(S \cap T) \le \max(v(S), v(T)) + \min(v(S), v(T))$$
$$= v(S) + v(T).$$

Thus, G is submodular. We now use the result in [3] that every reliability extension of a convex game is convex. A game is convex (supermodular) if  $v(S \cup T) + v(S \cap T) > v(S) + v(T)$ for all  $S, T \subseteq N$ . Their proof, with the directions of inequalities reversed, shows that reliability extensions of submodular games are submodular. Thus, reliability extensions of Max-Games are submodular, and hence subadditive.  $\blacksquare$ 

#### **THE CORE & THE CoS** 3.

With the results of Section 2.1 at our disposal, we are ready to analyze the core of Max-Games with failures. We have  $v(N) = \sum_{i=1}^{n} D(N_b^i) \cdot r_i w_i \leq \sum_{i=1}^{n} r_i w_i = \sum_{i=1}^{n} v(\{i\})$  due to Lemma 1. With at least two agents, the inequality becomes strict, so the value of the grand coalition is not sufficient to pay every agent at least its value. Thus:

THEOREM 1. The core of any reliability extension of a Max-Game with at least two agents is empty.

As discussed in Section 2, cooperation is clearly infeasible in Max-Games without failures. Failures present a tradeoff to high weight agents in cooperating with low weight agents, as the latter now add value to the coalition but also share the payoff. While Theorem 1 shows that this tradeoff does not help eliminate emptiness of the core, we show that it does reduce agents' resistance to cooperation, quantifying it using the Cost of Stability [2], which measures the minimal subsidy required to achieve stability (see Section 1.1).

Reliability extensions of Max-Games are subadditive (Lemma 2). Consider any subadditive game G = (N, v), and any payment vector **p**. If  $p_i \ge v(\{i\})$  for all  $i \in N$ , then  $p(C) = \sum_{i \in C} p_i \ge \sum_{i \in C} v(\{i\}) \ge v(C)$  for all  $C \subseteq N$ , where the last step is due to subadditivity. Thus, a subsidy  $\Delta$  ensures that the core is non-empty if and only if the increased value of the grand coalition is sufficient to give each agent at least its value, that is if  $v(N) + \Delta \geq \sum_{i \in N} v(\{i\})$ . Hence, the Cost of Stability  $CoS(G) = \sum_{i \in N} v(\{i\}) - v(N)$ . For any Max-Game  $G = (N, \mathbf{w})$  and its reliability exten-

sion  $G^{\mathbf{r}} = (N, \mathbf{w}, \mathbf{r}), CoS(G^{\mathbf{r}})$  would equal to

$$\sum_{i=1}^{n} v^{\mathbf{r}}(\{i\}) - v^{\mathbf{r}}(N) = \sum_{i=1}^{n} (1 - D(N_b^i)) \cdot r_i w_i \le \sum_{i=2}^{n} w_i.$$

The first step follows by Lemma 1 and since  $v^{\mathbf{r}}(\{i\}) = r_i w_i$ for every agent  $i \in N$ . The last step follows since the coefficient of  $w_1$  is zero and that of  $w_i$  is at most 1 for  $i \ge 2$ . Further,  $CoS(G) = \sum_{i=1}^{n} v(\{i\}) - v(N) = \sum_{i=2}^{n} w_i$ . Thus,

LEMMA 3. For any Max-Game G and its reliability extension  $G^{\mathbf{r}}$ , we have  $CoS(G^{\mathbf{r}}) \leq CoS(G)$ .

Thus, the presence of failures reduces the external subsidy required to induce agent cooperation, that is, it reduces agents' resistance to cooperation.

#### 4. POWER INDICES

We now show that two popular power indices, the Banzhaf index and the Shapley value, can be computed in polynomial time for any reliability extension of a Max-Game.

THEOREM 2. The Banzhaf index can be computed in polynomial time for any reliability extension of a Max-Game.

PROOF. Let  $G = (N, \mathbf{w})$  be a Max-Game and  $G^{\mathbf{r}} =$  $(N, \mathbf{w}, \mathbf{r})$  its reliability extension. Let v and  $v^{\mathbf{r}}$  be the characteristic functions of G and  $G^{\mathbf{r}}$ , respectively. Consider a survival process where each agent  $i \in N$  survives with probability  $r_i$  and fails with probability  $1-r_i$ . Let  $X \subseteq N$  denote the set of all surviving agents, and let  $X \sim 2^N$  denote that X is picked according to the survival process. The value of a coalition S in  $G^{\mathbf{r}}$  is the expected value of its surviving sub-coalition:

$$v^{\mathbf{r}}(S) = \mathbb{E}_{X \sim 2^N}[v(S \cap X)]. \tag{4}$$

Recall that the Banzhaf index  $\beta_i$  is the average marginal contribution of agent i to all coalitions. Alternatively, consider a *selection process* where we pick a coalition S by selecting every agent in N - i with an equal probability of 1/2. Let  $S \sim 2^{N-i}$  denote the coalition S picked this way (uniformly at random). Now,  $\beta_i$  is the *expected* marginal contribution of *i* to *S*, i.e.,  $\beta_i = \mathbb{E}_{S \sim 2^{N-i}} [v^{\mathbf{r}}(S+i) - v^{\mathbf{r}}(S)].$ Using (4) and the linearity of expectation,

$$\beta_i = \mathbb{E}_{S \sim 2^{N-i}, X \sim 2^N} [v((S+i) \cap X) - v(S \cap X)].$$

Consider the following cases.

- 1.  $i \notin X$  (i.e., *i* fails). Then,  $v((S+i) \cap X) v(S \cap X)$  is always 0.
- 2.  $i \in X$  (i.e., *i* survives). Then,  $v((S+i) \cap X) v(S \cap X)$  $X = v((S \cap X) + i) - v(S \cap X)$ . Consider three subcases: a) If  $S \cap X = \phi$ , then the marginal contribution is  $w_i$ ; b) If  $min(S \cap X) = j > i$ , then the marginal contribution is  $w_i - w_j$ ; c) If  $min(S \cap X) = j < i$ , then the marginal contribution is 0.

Let  $p_i$  and  $p_{ij}$  denote the probabilities of cases a) and b) respectively. Then,

$$\beta_{i} = p_{i}w_{i} + \sum_{j>i} p_{ij}(w_{i} - w_{j}) = \left(p_{i} + \sum_{j>i} p_{ij}\right)w_{i} - \sum_{j>i} p_{ij}w_{j}.$$
(5)

The coefficient of  $w_i$  is  $p_i + \sum_{j>i} p_{ij}$ , which is the total probability of case a) and b) (for any j > i). This occurs when agent *i* survives and no agent from  $\{1, \ldots, i-1\}$  is present in  $S \cap X$ . An agent  $t \in N - i$  is present in S with probability 1/2, and present in X independently with probability  $r_t$ , so present in  $S \cap X$  with probability  $r_t/2$ . Thus, the probability that t is not present in  $S \cap X$  is  $1 - r_t/2$ . Hence, the coefficient of  $w_i$  is  $r_i \cdot \prod_{t < i} (1 - r_t/2)$ . Consider an agent j where j > i. The coefficient of  $w_j$  in

Consider an agent j where j > i. The coefficient of  $w_j$  in  $\beta_i$  is  $-p_{ij}$ , where  $p_{ij}$  is the probability that i survives and  $min(S \cap X) = j$ . The latter requires that agents 1 through i-1 and i+1 through j-1 be absent from  $S \cap X$  and agent j be present in  $S \cap X$ . The total probability is  $r_i \cdot \left(\prod_{t < j, t \neq i} (1 - r_t/2)\right) \cdot r_j/2$ . Substituting these coefficients in (5) gives the explicit formula that allows computing the Banzhaf index in polynomial time.

Similarly, the Shapley value can also be computed in polynomial time for any Max-Game with failures. However, we were not able to use the elegant probabilistic technique due to the formula's dependence on the size of the coalition.

# THEOREM 3. The Shapley value can be computed in polynomial time for any reliability extension of a Max-Game.

PROOF. Let  $G = (N, \mathbf{w})$  be a Max-Game,  $G^{\mathbf{r}} = (N, \mathbf{w}, \mathbf{r})$  be its reliability extension, and v and  $v^{\mathbf{r}}$  be the characteristic functions of G and  $G^{\mathbf{r}}$ , respectively. Fix an agent  $i \in N$ . Substituting the formula for marginal contribution from Corollary 1 into the Shapley equation (2), we get

$$\phi_i = \frac{1}{n!} \sum_{S \subseteq N-i} |S|! (n - |S| - 1)! \cdot r_i \cdot D(S_b^i) \cdot \left(w_i - v^{\mathbf{r}}(S_a^i)\right).$$

After breaking the summation over  $S_b^i$  and  $S_a^i$ , and letting  $k = |S_b^i|$  and  $l = |S_a^i|$ , we get

$$\phi_{i} = \frac{r_{i}}{n!} \sum_{k=0}^{i-1} \sum_{l=0}^{n-i} \sum_{\substack{S_{b}^{i} \subseteq N_{b}^{i} \\ |S_{b}^{i}|=k}} \sum_{\substack{S_{a}^{i} \subseteq N_{a}^{i} \\ |S_{a}^{i}|=l}} (k+l)!(n-k-l-1)! \times D(S_{b}^{i}) \left(w_{i} - v^{\mathbf{r}}(S_{a}^{i})\right).$$

Note that  $N_b^i = \{1, \ldots, i-1\}$  and  $N_a^i = \{i+1, \ldots, n\}$ . For all  $l, t \leq n$ , define  $T(l, t) = \sum_{C \subseteq \{t, \ldots, n\}, |C|=l} v^{\mathbf{r}}(C)$ . Taking the summation over  $S_a^i$  inside,

$$\phi_{i} = \frac{r_{i}}{n!} \sum_{k=0}^{i-1} \sum_{l=0}^{n-i} \sum_{\substack{S_{b}^{i} \subseteq N_{b}^{i} \\ |S_{b}^{i}| = k}} \left[ (k+l)!(n-k-l-1)! \cdot D(S_{b}^{i}) \times \left( \binom{n-i}{l} w_{i} - T(l,i+1) \right) \right].$$

To sum over  $S_b^i$ , let  $F(k, a, b) = \sum_{C \subseteq \{a, \dots, b\}, |C|=k} D(C)$  for all  $a, b \in N$  such that a < b and all  $k \leq b - a + 1$ . Then,

the term  $D(S_b^i)$  is replaced by F(k, 1, i - 1). Now, taking the summation over l inside and rearranging terms gives the formula for the Shapley value in Table 1. Next, we simplify functions T and F. Let us examine the coefficient of  $w_j$  for  $j \ge t$  in T(l, t). Using Lemma 1, it is easy to verify that the coefficient is  $r_j \cdot \sum_{C \subseteq \{t, \dots, n\}, j \in C, |C| = l} D(C_b^j)$ . Let  $l' = |C_b^j|$ , so  $|C_a^j| = l - l' - 1$ . Note that  $l' \le \min(l - 1, j - t)$ . The coefficient simplifies to

$$r_{j} \cdot \sum_{l'=0}^{\min(l-1,j-t)} \left[ \sum_{\substack{C_{b}^{j} \subseteq \{t,\dots,j-1\} \\ |C_{b}^{j}| = l'}} D(C_{b}^{j}) \right] \cdot \left[ \sum_{\substack{C_{a}^{j} \subseteq \{j+1,\dots,n\} \\ |C_{a}^{j}| = l-l'-1}} 1 \right].$$

It is easy to verify that the summation over  $C_a^j$  is  $\binom{n-j}{l-l'-1}$ , and the summation over  $C_b^j$  is F(l', t, j-1). Substituting these gives the formula for T(l, t) in Table 1.

Next, note F(k, a, b) is exactly the  $k^{th}$  elementary symmetric polynomial [18] for the set  $\{1 - r_a, \ldots, 1 - r_b\}$ . The formula for F(k, a, b) given in Table 1 is known as Newton's identity [19]. Finally, observe that a direct implementation of the formulae in Table 1 finds the Shapley value in polynomial time.

# 5. INCENTIVES FOR SABOTAGE AND MANIPULATING RELIABILITY

In some domains, an agent may sabotage others or influence the failure probabilities of itself and its peers. For example, consider an agent that controls the time of performing a task at hand; a time slot may be good for some agents and bad for others, affecting agents' failure probabilities. The time slot should be chosen optimally to maximize the value of the grand coalition. If the agents decide to distribute the gains using the Shapley value, could the controlling agent benefit from choosing a different time slot, thus manipulating reliabilities of its peers? We answer such questions in the context of Max-Games.

Consider a reliability extension of a Max-Game with two agents,  $w_1 = w_2 = 1$ , and  $r_1 = r_2 = 0.5$ . The grand coalition's value is v(N) = 0.75, and the Shapley values are  $\phi_1 = \phi_2 = 0.375$ . Suppose agent 1 reduces the reliability of agent 2, changing  $r_2$  to 0.1. The grand coalition's value drops to v'(N) = 0.55, and the Shapley values change to  $\phi'_1 = 0.475$ ,  $\phi'_2 = 0.075$ . Thus, the sabotage is beneficial to agent 1, even though the grand coalition's value drops.

Such manipulations were not previously studied in cooperative games. We analyze them in reliability extensions of Max-Games. Consider any Max-Game with failures  $G^{\mathbf{r}} = (N, \mathbf{w}, \mathbf{r})$ . Fix agents  $i, j \in N$ . Consider a sabotage where agent *i* decreases the reliability of agent *j*. Due to Corollary 1, the marginal contribution of *i* to a coalition S is  $v^{\mathbf{r}}(S + i) - v^{\mathbf{r}}(S) = r_i \cdot D(S_b^i) \cdot (w_i - v^{\mathbf{r}}(S_a^i))$ . Now,  $D(S_b^i)$ , the probability that every agent in  $S_b^i$  fails, can only increase by the sabotage, and the value  $v^{\mathbf{r}}(S_a^i)$  can only decrease by the sabotage.<sup>1</sup> Thus, sabotage can only increase the marginal contribution of agent *i* to every coalition *S*. Also, due to Corollary 1, an agent's Shapley value is directly proportional to its own reliability. Thus, we have:

<sup>&</sup>lt;sup>1</sup>It is easy to show that the value of any coalition is monotonically increasing in the reliability of every agent in it.

$$\begin{split} \phi_i &= \frac{r_i}{n!} \cdot \sum_{k=1}^{i-1} \left[ F(k,1,i-1) \cdot \sum_{l=0}^{n-i} \left( (k+l)! \cdot (n-k-l-1)! \cdot \left[ \binom{n-i}{l} \cdot w_i - T(l,i+1) \right] \right) \right] \\ T(l,t) &= \sum_{j=t}^n \left[ r_j \cdot w_j \cdot \sum_{l'=0}^{\min(l-1,j-t)} \left( \binom{n-j}{l-l'-1} \cdot F(l',t,j-1) \right) \right], \quad T(0,t) = 0, \ \forall t \\ F(k,a,b) &= \frac{1}{k} \cdot \sum_{t=1}^k \left[ (-1)^{t-1} \cdot F(k-t,a,b) \cdot P(t,a,b) \right], \quad F(0,a,b) = 1, \ \forall a,b \\ P(k,a,b) &= \sum_{j=a}^b (1-r_j)^k \end{split}$$

Table 1: Shapley value in reliability extensions of Max-Games

THEOREM 4. In any reliability extension of a Max-Game, every agent is better off, in terms of its Shapley value, by increasing its own, and decreasing everyone else's reliability.

Now consider a Max-Game where failures are caused by a common factor, which applies equally to all agents, so all failures probabilities are equal. For example, all agents may be similarly affected by the choice of time slot, so the agent that controls the time slot can choose the uniform failure probability, but cannot affect any agent's failure probability individually. If agent *i* increases the common failure probability, the rise in others' failure probabilities is beneficial to *i*, but the rise in its own failure probability is harmful. Could increasing the common failure probability, and thus harming the entire project's reward, be beneficial to *i* overall? We answer this negatively.

THEOREM 5. In any reliability extension of a Max-Game where all failure probabilities are equal, no agent is better off, in terms of its Shapley value, by increasing the common failure probability.

To show this, consider any Max-Game  $G = (N, \mathbf{w})$ . Let  $\phi_i(x)$  be the Shapley value of agent *i* in the reliability extension  $G^{\mathbf{r}} = (N, \mathbf{w}, \mathbf{r})$ , where  $r_j = x$  for all  $j \in N$ . Then, we want to show that for every agent *i* and all p, p' such that  $0 \leq p < p' \leq 1$ ,  $\phi_i(p) < \phi_i(p')$ . One way to obtain  $\phi_i(x)$  is to substitute all  $r_i$ 's equal to x in the Shapley value formula for the non-uniform case given in Table 1. However, we can derive the following simple formula through the probabilistic technique used in the proof of Theorem 2.

LEMMA 4. The value of  $\phi_i(p)$  is given by

$$w_i \left[ \frac{1 - (1 - p)^i}{i} \right] - \sum_{j=i+1}^n w_j \left[ \frac{1 - (1 - p)^{j-1}}{j-1} - \frac{1 - (1 - p)^j}{j} \right].$$

PROOF. Fix any  $p \in [0, 1]$ . We denote  $\phi_i(p)$  by  $\phi_i$  in the remainder of the proof. The Shapley value of agent i is its average marginal contribution to its predecessors over all agent permutations. Let  $S_n$  be the set of all permutations of  $\{1, \ldots, n\}$  and let  $\pi \sim S_n$  denote that  $\pi$  is chosen uniformly at random from  $S_n$ . Now, the Shapley value of agent i is the *expected* marginal contribution of i to  $\Gamma_i^{\pi}$  when  $\pi \sim S_n$ , i.e.,  $\phi_i = \mathbb{E}_{\pi \sim S_n}[v^{\mathbf{r}}(\Gamma_i^{\pi} + i) - v^{\mathbf{r}}(\Gamma_i^{\pi})]$ . As in the proof of Theorem 2, let X denote the set of surviving agents. Using (4), we have  $\phi_i = \mathbb{E}_{\pi \sim S_n, X \sim 2^N} \left[ v \left( (\Gamma_i^{\pi} + i) \cap X \right) - v (\Gamma_i^{\pi} \cap X) \right]$ . We now examine the same cases as considered in the proof of Theorem 2, but the coalition S is now replaced by  $\Gamma_i^{\pi}$ . Let  $p_i$  and  $p_{ij}$  be the corresponding probabilities as defined in the proof of Theorem 2. We have  $\phi_i = \left(p_i + \sum_{j=i+1}^n p_{ij}\right) \cdot w_i - \sum_{j=i+1}^n p_{ij} \cdot w_j$ .

 $w_i - \sum_{j=i+1}^n p_{ij} \cdot w_j$ . The coefficient of  $w_i$  is  $p_i + \sum_{j>i} p_{ij}$ , which is the total probability of case a) and case b) (for any j > i). It is the probability that agent *i* survives and no agent from  $\{1, \ldots, i-1\}$  is present in  $\Gamma_i^{\pi} \cap X$ . That is, all agents from  $\{1, \ldots, i-1\}$  that are before agent *i* in  $\pi$  fail. Let  $\sigma$  be the permutation denoting the order of agents  $\{1, \ldots, i\}$  in  $\pi$ .<sup>2</sup> Let  $F_k^{\sigma}$  denote the event that agent  $\sigma^{-1}(k)$  fails. By independence, the probability that agent *i* survives and agents from  $\{1, \ldots, i-1\}$  before agent *i* in  $\pi$  fail is:

$$p \cdot \sum_{k=1}^{i} \Pr[(\sigma(i) = k) \land F_{1}^{\sigma} \land \dots \land F_{k-1}^{\sigma}]$$
$$= p \cdot \sum_{k=1}^{i} \frac{1}{i} \cdot (1-p)^{k-1} = \frac{1 - (1-p)^{i}}{i}$$

This is because  $\sigma(i)$  takes each value  $1 \le k \le i$  with probability 1/i, and survival of agents in positions 1 to k - 1 in  $\sigma$  is independent of agent *i*'s position.

Next, for any j > i, the coefficient of  $w_j$  is  $-p_{ij}$ , where  $p_{ij}$  is the probability of case b) with  $min(\Gamma_i^{\pi} \cap X) = j$ . Let  $\tau$  be the sub-permutation of  $\pi$  over agents  $\{1, \ldots, j\}$ . Then  $\tau(j) < \tau(i)$ , and every agent from  $\{1, \ldots, j-1\}$  before *i* in  $\tau$  must fail (or  $min(\Gamma_i^{\pi} \cap X) \neq j$ ). Also, both agents *i* and *j* must survive. Let  $F_k^{\tau}$  be the event that agent  $\tau^{-1}(k)$  fails. Denoting positions of *j* and *i* by  $k_1$  and  $k_2$ ,

$$p_{ij} = \sum_{k_1=1}^{j-1} \sum_{k_2=k_1+1}^{j} \Pr\left[ (i, j \text{ survive}) \land (\tau(j) = k_1) \land (\tau(i) = k_2) \land \bigwedge_{\substack{t=1\\s.t.t \neq k_1}}^{k_2-1} F_t^{\tau} \right]$$
$$= p^2 \cdot \sum_{k_1=1}^{j-1} \sum_{k_2=k_1+1}^{j} \left( \Pr[\tau(j) = k_1] \times \Pr[\tau(i) = k_2 | \tau(j) = k_1] \cdot (1-p)^{k_2-2} \right)$$

<sup>2</sup>Formally,  $\sigma(l) = |\{t \in \{1, ..., i\} | \pi(t) \le \pi(l)\}|$ , for  $l \le i$ .

$$= \frac{p^2}{j \cdot (j-1)} \cdot \sum_{k_1=1}^{j-1} \sum_{k_2=k_1+1}^{j} (1-p)^{k_2-2}$$
$$= \frac{p^2}{j \cdot (j-1)} \cdot \sum_{k_2=2}^{j} (k_2-1) \cdot (1-p)^{k_2-2}.$$
 (6)

The second transition follows since agent failures are independent each other and of agent positions. So given  $\tau(j)$ , there are j - 1 equiprobable values for  $\tau(i)$ . The fourth transition follows by reversing the order of summations and subsequent simplification. Multiplying by 1 - p:

$$(1-p)p_{ij} = \frac{p^2}{j(j-1)} \sum_{k_2=2}^{j} (k_2-1)(1-p)^{k_2-1}$$
$$= \frac{p^2}{j(j-1)} \sum_{k_2=3}^{j+1} (k_2-2)(1-p)^{k_2-2}.$$
(7)

Now, subtracting (7) from (6), and then dividing by p:

$$p_{ij} = \frac{p}{j(j-1)} \left( 1 - (j-1)(1-p)^{j-1} + \sum_{k_2=3}^{j} (1-p)^{k_2-2} \right)^{j}$$
$$= \frac{p}{j(j-1)} \cdot \left( -(j-1)(1-p)^{j-1} + \sum_{t=0}^{j-2} (1-p)^t \right)^{j}$$
$$= \frac{1 - (1-p)^{j-1}}{j-1} - \frac{1 - (1-p)^j}{j}.$$

Substituting these coefficients in the Shapley formula completes the proof.  $\blacksquare$ 

Note that Lemma 4 gives a simple formula for the Shapley value in Max-Games without failures (p = 1), namely,  $\phi_i = w_i/i - \sum_{j=i+1}^n w_j/(j \cdot (j - 1))$ . Again, it is easy to verify that the Shapley value thus obtained is independent of the tie-breaking used among the equal weight agents. We now prove Theorem 5.

PROOF. To prove that  $\phi_i(p)$  is strictly increasing for  $p \in [0, 1]$ , we show that its derivative  $\phi'_i(p) > 0$  for  $p \in (0, 1)$ . By Lemma 4,

$$\phi'_i(p) = w_i \cdot (1-p)^{i-1} - \sum_{j=i+1}^n w_j \cdot \left( (1-p)^{j-2} - (1-p)^{j-1} \right).$$

Putting  $w_j \leq w_i$  for j > i (agents are sorted by weight) and simplifying, we get

$$\phi'_i(p) \ge w_i\left((1-p)^{i-1} - p\sum_{j>i}(1-p)^{j-2}\right) = (1-p)^{n-1} > 0,$$

which is the required result.  $\blacksquare$ 

#### 6. RELATED WORK

Agent failures were investigated in non-cooperative games, such as congestion games [23, 21] and in network domains [8, 17]. In contrast, we focus on *cooperation* between selfish agents, using a cooperative game [22] and its reliability extension [4]. Cooperative games were used to study negotiation, reward sharing and team formation [25, 29]. We use solutions such as the core [15], the Shapley value [28], and the Banzhaf index [9]. We used the Cost of Stability to quantify resistance to cooperation, similarly to earlier work dealing with other games [24, 20, 6].

Our model bears some similarity to Combinatorial Agency [1, 7] and all-pay auctions [10]. In Combinatorial Agency, a principal has to reward agents based only on the final outcome without observing their exerted efforts, while in Max-Games with agent failures, one rewards the agents ex-ante, without observing if they participated or not. In all-pay auctions, the principal solicits multiple submissions from agents, and incentivizes sincere efforts by promising a reward to the agent whose submission was of the highest quality. Hence, they are similar to Max-Games in that they use redundancy to improve the overall quality, but the key difference is that in Max-Games we reward all agents rather than just one in order to incentivize a large number of participants. The model of crowdsourcing contests proposed in [12] is fundamentally different than ours (maximizing efficiency in a non-cooperative game), but shares the idea of rewarding agents besides the one with the highest value. Max-Games also differ from both these models since in our case agent failures are not the result of strategic decisions of the agents about how much effort to exert; rather, they stem from the potential of failures due to *external* reasons.

Finally, agent weights in Max-Games are reminiscent of weighted voting games (WVGs). In WVGs, the value of a coalition is an additive aggregation of agent weights. In contrast, Max-Games model situations where a coalition's value depends on the *maximal* weight in the coalition. Max-Games are also reminiscent of skill games [5] since agent weights in Max-Games represent agents' ability levels, but contrary to skill games, Max-Games do not rely on a set cover combinatorial structure. Also, the polynomial time algorithms presented in this paper for computing power indices in Max-Games (even under failures) sharply contrast the  $\mathcal{NP}$ -hardness or even  $\#\mathcal{P}$ -hardness of computing power indices in WVGs [13] and skill games [5], respectively.

### 7. DISCUSSION

We proposed Max-Games where the value of a coalition is determined by the maximal quality of the agents in it, and examined the impact of agent failures in such games. We analyzed solutions to the game, and provided polynomialtime algorithms for computing power indices. An important conceptual contribution of this paper is the analysis of incentives introduced by agent failures in cooperative games, where an agent wants to sabotage its peers by increasing their failure probabilities. We initiated the study of such incentives, analyzed them for Max-Games, and provided a positive result for the case of uniform failure probabilities.

We showed that the core of a Max-Game is always empty, with or without failures. Several relaxations of the core have been proposed in the literature. The  $\epsilon$ -core is the set of "reasonably-stable" payoff allocations [27], and the nucleolus is the "most-stable" payoff allocation [26]. When payments are non-negative, it can be shown that checking emptiness of the  $\epsilon$ -core and finding the nucleolus of Max-Games without failures can be done in polynomial time. However, we were unable to settle these questions for Max-Games with failures. We showed that *resistance* to cooperation diminishes when failures are introduced, using the Cost of Stability. Agent failures have been shown to increase stability in other games using qualitative measures such as non-emptiness of the core, and quantitative measures such as the least core value and the Cost of Stability [4, 3, 6]. It would be interesting to perform a similar detailed analysis for Max-Games.

Our model is based on a *revenue sharing* cooperative game, where the agents share the gains from cooperating. A natural extension of this work is examining *cost sharing* cooperative games, where each agent must accomplish a goal associated with a certain cost, and where agents may collaborate so as to share the cost. A cost sharing model in the same spirit as our Max-Game is one where each agent has a weight (reflecting the cost of achieving the goal), and the cost of a coalition is the minimal weight of any agent in the coalition. More generally, the cost of a coalition can be any reasonable function of the weights of the agents in the coalition. It would be interesting to study this costsharing game, and to find the relation between its solutions and solutions to our Max-Game.

Finally, we remark that our analysis of incentives for sabotaging peers, i.e., perturbing their failure probabilities, is presented for Max-Games (and its restricted forms), but its scope extends to all cooperative games. It would be interesting to study these incentives in other important classes of cooperative games such as weighted voting games. Consider the "no sabotage" result of Theorem 5 for restricted Max-Games. Such a result also depends on the payoff scheme used, which is the Shapley value in our case. So, we ask: For what other classes of cooperative games and under what payoff schemes can we obtain "no sabotage"? Also, in games that do have incentives for sabotage, which agents would a given agent wish to make more (or less) reliable? When such sabotage is *costly* and the saboteur has a limited budget, how can it find the optimal manipulation?

One can imagine the "no sabotage" property as a new axiom for payoff division, and study its interactions with the four axioms of the Shapley value. We believe that such interactions may give rise to fundamentally new payoff schemes for cooperative games which are both fair and robust.

#### 8. **REFERENCES**

- M. Babaioff, M. Feldman, and N. Nisan. Combinatorial agency. In *Proc. of 7th EC*, pages 18–28, 2006.
- [2] Y. Bachrach, E. Elkind, R. Meir, D. Pasechnik, M. Zuckerman, J. Rothe, and J. S. Rosenschein. The cost of stability in coalitional games. In *Proc. of 2nd SAGT*, pages 122–134, 2009.
- [3] Y. Bachrach, I. Kash, and N. Shah. Agent failures in totally balanced games and convex games. In *Proc. of* 8th WINE, pages 15–29, 2012.
- Y. Bachrach, R. Meir, M. Feldman, and M. Tennenholtz. Solving cooperative reliability games. In *Proc. of 27th UAI*, pages 27–34, 2011.
- [5] Y. Bachrach and J. S. Rosenschein. Coalitional skill games. In *Proc. of 7th AAMAS*, pages 1023–1030, 2008.
- [6] Y. Bachrach and N. Shah. Reliability weighted voting games. In Proc. of 6th SAGT, 2013.
- [7] Y. Bachrach, M. Zuckerman, and J. Rosenschein. Effort Games and the Price of Myopia. *Mathematical Logic Quarterly*, 55(4):377–396, 2009.
- [8] V. Bala and S. Goyal. A strategic analysis of network reliability. *Review of Economic Design*, 5:205–228, 2000.

- [9] J. F. Banzhaf. Weighted voting doesn't work: a mathematical analysis. *Rutgers Law Review*, 19:317–343, 1965.
- [10] M. R. Baye, D. Kovenock, and C. G. De Vries. The all-pay auction with complete information. *Economic Theory*, 8(2):291–305, 1996.
- [11] J. Bennett, C. Elkan, B. Liu, P. Smyth, and D. Tikk. KDD cup and workshop 2007. SIGKDD Explorations, 9(2):51–52, 2007.
- [12] R. Cavallo and S. Jain. Efficient crowdsourcing contests. In *Proc. of 11th AAMAS*, pages 677–686, 2012.
- [13] E. Elkind, L. A. Goldberg, P. W. Goldberg, and M. Wooldridge. Computational complexity of weighted threshold games. In *Proc. of 22nd AAAI*, pages 718–723, 2007.
- [14] X. Gao, Y. Bachrach, P. Key, and T. Graepel. Quality expectation-variance tradeoffs in crowdsourcing contests. In *Proc. of 26th AAAI*, pages 38–44, 2012.
- [15] D. Gillies. Some Theorems on n-Person Games. PhD thesis, Princeton U., 1953.
- [16] J. Howe. The rise of crowdsourcing. Wired magazine, 14(6):1–4, 2006.
- [17] R. Kannan, S. Sarangi, and S. Iyengar. A simple model for reliable query reporting in sensor networks. In *Information Fusion*, volume 2, pages 1070 – 1075, 2002.
- [18] I. Macdonald. Symmetric functions and Hall polynomials. New York, 1995.
- [19] D. Mead. Newton's identities. The American mathematical monthly, 99(8):749–751, 1992.
- [20] R. Meir, J. Rosenschein, and E. Malizia. Subsidies, stability, and restricted cooperation in coalitional games. In *Proc. of 22nd IJCAI*, pages 301–306, 2011.
- [21] R. Meir, M. Tennenholtz, Y. Bachrach, and P. Key. Congestion games with agent failures. In Proc. of 26th AAAI, pages 1401–1407, 2012.
- [22] M. Osborne and A. Rubinstein. A course in game theory. The MIT press, 1994.
- [23] M. Penn, M. Polukarov, and M. Tennenholtz. Congestion games with failures. *Discrete Applied Mathematics*, 159(15):1508–1525, 2011.
- [24] E. Resnick, Y. Bachrach, R. Meir, and J. Rosenschein. The cost of stability in network flow games. In *Proc. of 34th MFCS*, pages 636–650, 2009.
- [25] J. Rosenschein and G. Zlotkin. Rules of encounter: designing conventions for automated negotiation among computers. MIT press, 1994.
- [26] D. Schmeidler. The nucleolus of a characteristic function game. SIAM J. Appl. Math., 17(6):1163–1170, 1969.
- [27] L. Shapley and M. Shubik. Quasi-cores in a monetary economy with nonconvex preferences. *Econometrica: Journal of the Econometric Society*, pages 805–827, 1966.
- [28] L. S. Shapley. A value for n-person games. Contrib. to the Theory of Games, pages 31–40, 1953.
- [29] Y. Shoham and K. Leyton-Brown. Multiagent systems: Algorithmic, game-theoretic, and logical foundations. Cambridge University Press, 2008.