Disarmament Games with Resources

Yuan Deng, Vincent Conitzer

Department of Computer Science
Duke University
Durham, NC 27708, USA
{ericdy,conitzer}@cs.duke.edu

Abstract

A paper by Deng and Conitzer in AAAI'17 introduces *disarmament games*, in which players alternatingly commit *not* to play certain pure strategies. However, in practice disarmament usually does not consist in removing a *strategy*, but rather in removing a *resource* (and doing so rules out all the strategies in which that resource is used simultaneously). In this paper, we introduce a model of disarmament games in which resources, rather than strategies, are removed. We prove NP-completeness of several formulations of the problem of achieving desirable outcomes via disarmament. We then study the case where resources can be *fractionally* removed, and prove a result analogous to the folk theorem that all desirable outcomes can be achieved. We show that we can approximately achieve any desirable outcome in a polynomial number of rounds, though determining whether a given outcome can be obtained in a given number of rounds remains NP-complete.

Introduction

Computing game-theoretic solutions has long been a topic of interest to AI researchers and other computer scientists. In earlier days, the main focus was on computing Nash equilibria (Gilboa and Zemel 1989; Porter, Nudelman, and Shoham 2008; Conitzer and Sandholm 2008; Daskalakis, Goldberg, and Papadimitriou 2009; Chen, Deng, and Teng 2009). While computing Nash equilibria remains a very active topic of research, much of the recent work on computing game-theoretic solutions done in the AI community has focused on computing optimal strategies to *commit* to, also referred to as *Stackelberg strategies* (Conitzer and Sandholm 2006; von Stengel and Zamir 2010). Algorithms for computing these solutions have been deployed in a variety of real-world domains (Tambe 2011). All this work concerns models in which one player commits, and then another responds to the commitment. A recent exception is a paper by Deng and Conitzer (2017), who show that sometimes, much more can be achieved by allowing players to go back and forth, alternatingly committing *not* to play certain pure strategies. They call the resulting games *disarmament games*.

In these disarmament games, there is a *disarmament phase* before the players play the game. In a step of this phase, the current player can remove any of his pure strategies in the normal-form game. Once players stop disarming, the players play the normal-form game consisting of the remaining pure strategies. This results in a larger (extensive-form) game that includes the disarmament phase, and the goal is to identify desirable equilibria of this larger game, which generally result in much better outcomes than the equilibria of the game without disarmament.[1] That is, we try to have players disarm so as to achieve a better outcome, but we must ensure that players are *incentivized* to take these disarmament steps.

However, in real disarmament games, the removal of *individual pure strategies* would often be hard to achieve. It is often possible to remove a *resource* (say, a particular weapon), but doing so generally eliminates *many* strategies, namely all those that involved the use of that resource. Moreover, removing resources does not allow one to remove the pure strategy of *not using any resources at all*. Finally, even in an environment where it is somehow possible to remove individual pure strategies (i.e., subsets of resources), the number of pure strategies is exponential. Moreover, Deng and Conitzer (2017) only consider Nash equilibrium as the solution concept, which may consist of incredible threats in subgames of the extensive-form disarmament games. However, it is hard to generalize to subgame perfect equilibrium when considering normal-form games since the computation of Nash equilibrium in a normal-form game is already hard.

Motivated by these observations, in this paper, we directly model the removal of resources instead of pure strategies. To do so, we define *multi-resource* games, in which each player has some resources, the use of which will impact the players differently. We then consider a disarmament model for these games and investigate whether and how we can reach desirable outcomes for several variants (requiring subgame perfection or not and allowing fractional removals or not).

Example 1. *Suppose that there are two players,* 0 *and* 1. *Each of them has two resources,* x *and* y. *Using* x *results in* 1 *utility for the player who plays it and* -2 *utility for the opponent; using* y *results in* 0 *utility for the player who plays*

---

[1]There will also always be an undesirable equilibrium of the disarmament game where nobody disarms, but it seems reasonable to believe that better equilibria will be played in practice, if they can be identified.

it and -2 utility for the opponent. Without disarmament, the Pareto optimal equilibrium of this game is that both players use only x, resulting in -1 utility for each.

However, with disarmament, it is possible to remove all resources, so that both players end up with utility 0, a clear Pareto improvement. To achieve this, we can use the following disarmament schedule. First, player 0 removes x; then, player 1 removes x and y; and finally, player 0 removes y. This disarmament schedule can be sustained in an equilibrium of the larger disarmament game, because if player 1 decides to deviate from the schedule and not disarm, then 0 can play resource y, resulting in a utility of at most -1 for player 1, lower than what 1 would obtain for not deviating.

Definitions

We first introduce multi-resource games and then define disarmament games based on the multi-resource games. We only consider multi-resource games in which there are no interactions between the resources, i.e., their effects on the players' utilities are additive. This restriction only strengthens the hardness results that we prove, because hardness results are stronger when they are proved in a more restricted setting. The positive results for fractional disarmament that we prove later in the paper can be generalized to hold under a much weaker assumption than additivity, but we do not do so here for the sake of readability.

Definition 1. *A 2-player multi-resource game is defined by $G = \langle R_0, R_1, v_0, v_1 \rangle$. For each player $b \in \{0, 1\}$, his set of resources is R_b and his strategy set is the power set 2^{R_b} of his resource set R_b. His valuation function is $v_b : R_0 \cup R_1 \to \mathbb{R}$, where $v_b(s)$ denotes player b's valuation for resource $s \in R_0 \cup R_1$. The utility $u_b(S_0, S_1)$ of player b when player 0 plays $S_0 \in 2^{R_0}$ and 1 plays $S_1 \in 2^{R_1}$ is*

$$u_b(S_0, S_1) = \sum_{s \in S_0 \cup S_1} v_b(s).$$

In multi-resource games, we have the following simple characterization of Nash equilibria, which allows us to quickly compute a Nash equilibrium or even an optimal Nash equilibrium. For convenience, let $P_b = \{s \in R_0 \cup R_1 \mid v_b(s) > 0\}$ be the set of resources that give player b positive utility and let $N_b = \{s \in R_0 \cup R_1 \mid v_b(s) < 0\}$ be the set of resources that give player b negative utility.

Proposition 1. *In a multi-resource game $G = \langle R_0, R_1, v_0, v_1 \rangle$, the strategy profile (S_0, S_1) forms a Nash equilibrium if and only if $R_b \cap P_b \subseteq S_b \subseteq R_b \setminus N_b$ for $b \in \{0, 1\}$.*

The game induced by $T_0 \subseteq R_0$ and $T_1 \subseteq R_1$ is the two-player multi-resource game $G_{T_0, T_1} = \langle T_0, T_1, v_0, v_1 \rangle$, where v_0 and v_1 are restricted to $T_0 \cup T_1$. As usual, we use $-b$ to denote the player other than b.

We now define the disarmament game $G_D(G)$ on top of this multi-resource game. The remaining definitions in this section very closely mirror (and are in some cases identical to) those in (Deng and Conitzer 2017), and so should be seen as background rather than as a new contribution. The main difference is that in that earlier work, specific strategies of the normal-form game are removed, whereas in this

work, resources are removed (where the removal of a single resource generally results in the elimination of many strategies). We also consider subgame-perfect Nash equilibrium in this paper.

The disarmament game consists of two stages, a disarmament stage and a game play stage. During the disarmament stage, players alternatingly remove nonempty sets of resources from R_0 and R_1. In the game play stage, triggered when a player removes nothing, they play whatever multi-resource game remains. We then consider the resulting game as a perfect information extensive-form game. To be more precise, we consider a game DAG rather than a game tree, because multiple sequences of disarmaments may lead to the same state. Also, we associate terminal nodes of this DAG with multi-resource games, rather than (directly) with payoffs. This DAG still has exponential size, however, so we never represent it explicitly. All of this is consistent with prior work (Deng and Conitzer 2017).

Definition 2 (Disarmament Game). *Given a multi-resource game $G = \langle R_0, R_1, v_0, v_1 \rangle$, the disarmament game $G_D(G)$ is defined as an extensive-form game by the following.*

- *The set of disarmament actions A: $\{X_0 \mid X_0 \subsetneq R_0\} \cup \{X_1 \mid X_1 \subsetneq R_1\} \cup \{Play\}$, where X_b denotes the set of resources to keep and Play denotes ending the disarmament stage.*
- *The set of non-terminal nodes H: $\{[T_0, T_1, b] \mid T_0 \subseteq R_0, T_1 \subseteq R_1, b \in \{0, 1\}\}$, where T_0 and T_1 denote the remaining resources and b denotes the player to move.*
- *The set of terminal nodes Z: $\{[T_0, T_1] \mid T_0 \subseteq R_0, T_1 \subseteq R_1\}$.*
- *The player selection function $\rho : H \to \{0, 1\}$: $\rho([T_0, T_1, b]) = b$.*
- *The available-actions function $\chi : H \to 2^A$, where $\chi([T_0, T_1, b]) = \{X_b \mid X_b \subsetneq T_b\} \cup \{Play\}$. (Playing X_b corresponds to keeping those resources.)*
- *The successor function $\gamma : H \times A \to H \cup Z$:*
  - *$\gamma([T_0, T_1, 0], X_0) = [X_0, T_1, 1]$;*
  - *$\gamma([T_0, T_1, 1], X_1) = [T_0, X_1, 0]$;*
  - *$\gamma([T_0, T_1, b], Play) = [T_0, T_1]$.*
- *The root of the game is $root = [R_0, R_1, 0]$.*

In the terminal nodes $z = [T_0, T_1]$, player 0 and 1 play the multi-resource game G_{T_0, T_1}, resulting in their final utilities.

Definition 3 (Strategy in G_D). *A strategy $\sigma_b = (\alpha, \beta)$ for G_D consists of disarmament strategy $\alpha \in \prod_{h \in H \mid \rho(h) = b} \chi(h)$ for non-terminal nodes and play strategy $\beta \in \prod_{[T_0, T_1] \in Z} 2^{T_b}$ (i.e., subsets of remaining resources) for terminal nodes.*

Note that we restrict our attention to deterministic behavior during the disarmament stage (in which there is perfect information).

Definition 4 (On-path history & outcome). *For a strategy profile (σ_0, σ_1) with $\sigma_0 = (\alpha_0, \beta_0)$ and $\sigma_1 = (\alpha_1, \beta_1)$, denote the on-path history by $P = (h_0 = root, h_1, \cdots, h_K, h_{K+1} = z = [T_0, T_1])$ where for all i, $\gamma(h_i, \alpha_{\rho(h_i)}(h_i)) = h_{i+1}$. The outcome of the strategy profile (σ_0, σ_1) is $(\beta_0(z), \beta_1(z))$.*

We say an on-path history has length K if it contains K non-terminal nodes, excluding the root. In a slight abuse of notation, let $u_b(o)$ be player b's utility for outcome o.

We are interested in two equilibrium notions, namely Nash equilibrium (NE) and sub-game perfect Nash equilibrium (SPNE). A strategy profile (σ_0, σ_1) forms an NE if and only if no player can increase his utility by deviating to another strategy; and a strategy profile (σ_0, σ_1) forms an SPNE if and only if no player can increase his utility starting from *any* node by deviating to another strategy.

Deng and Conitzer (2017) shows that on-path histories of a strategy profile can serve as certificates to check whether a strategy profile forms a Nash equilibrium or not, placing the equilibrium computation in NP, if one can compute the *security level* for each player at each node of the game.

Definition 5 (Security level). *The security level sec_b for player b in a two-player multi-resource game $G = \langle T_0, T_1, v_0, v_1 \rangle$ is the utility that player b can guarantee himself no matter how the other player plays. Formally,*

$$sec_b(G_{T_0,T_1}) = \max_{\beta_b \in 2^{T_b}} \min_{\beta_{-b} \in 2^{T_{-b}}} u_b(\beta_b, \beta_{-b})$$

Proposition 2. *In a multi-resource game $G = \langle T_0, T_1, v_0, v_1 \rangle$, the security level for player b is*

$$sec_b(G_{T_0,T_1}) = \sum_{s \in P_b \cap T_b} v_b(s) + \sum_{s \in N_b \cap T_{-b}} v_b(s)$$

Lemma 1 (Deng and Conitzer (2017)). *P is an on-path history induced by a Nash equilibrium strategy profile with outcome o if and only if o is an equilibrium of the terminal node in P and for each non-terminal node $[T_0, T_1, b] \in P$, $sec_b(G_{T_0,T_1}) \leq u_b(o)$.*

Computing One Equilibrium

In disarmament games with resources, it is easy to compute one SPNE of the disarmament game, which implies that it is also easy to compute one NE.

Theorem 1. *In a multi-resource game $G = \langle R_0, R_1, v_0, v_1 \rangle$, given a Nash equilibrium strategy profile (S_0, S_1) (without disarmament) with utility $u_0(S_0, S_1) = u_0^*$ and $u_1(S_0, S_1) = u_1^*$, we can efficiently construct a subgame perfect Nash equilibrium of the corresponding disarmament game $G_D(G)$ with the same utility (u_0^*, u_1^*).*

Proof. Because (S_0, S_1) is a Nash equilibrium, by Proposition 1 we have $R_b \cap P_b \subseteq S_b \subseteq R_b \setminus N_b$ for $b \in \{0, 1\}$. Consider the following strategy profile: for each player b,

- for every non-terminal node, choose Play;
- for every terminal node $[T_0, T_1]$, choose $T_b \cap S_b$.

First, note that the players play an NE in every terminal node. As for the non-terminal nodes $[T_0, T_1, b]$, a deviation must change the state to $[X_b, T_{-b}, 1 - b]$ with $X_b \subsetneq T_b$. Next, $-b$ will choose Play, and subsequently $-b$ will play $T_{-b} \cap S_{-b}$. But this is exactly what $-b$ would have played if b had not deviated; moreover, without deviating, b would have best-responded to $T_{-b} \cap S_{-b}$ from within its resources T_0. Since $X_b \subsetneq T_b$, b can do no better after deviating. $\square$

Computing Equilibrium with Specific Outcome

Therefore, we are interested in the computation of an equilibrium that results in a particular outcome. We now define the two computational problems:

Definition 6 (RESOURCE-DISARM$^\mathcal{C}$ problem). *In RESOURCE-DISARM$^\mathcal{C}$ (where $\mathcal{C} \in \{NE, SPNE\}$) given a disarmament game G_D and an outcome $o^* = (\beta_0^*, \beta_1^*)$, the objective is to determine whether there exists an equilibrium (σ_0^*, σ_1^*) of type $\mathcal{C}$ such that the outcome is o^*.*

We also consider a variation of the RESOURCE-DISARM problem called K-RESOURCE-DISARM.

Definition 7 (K-RESOURCE-DISARM$^\mathcal{C}$ problem). *In K-RESOURCE-DISARM$^\mathcal{C}$, given a disarmament game G_D and an outcome $o^* = (\beta_0^*, \beta_1^*)$, the objective is to determine whether there exists an equilibrium (σ_0^*, σ_1^*) of type $\mathcal{C}$ such that the outcome is o^* and the length of its induced on-path history is at most K.*

RESOURCE-DISARM$^{\textbf{NE}}$

For an on-path history with length K, the total number of non-terminal nodes, excluding the root, is K. For $K \leq 2$, K-RESOURCE-DISARM^{NE} is in P: since both players have only one step to remove resources, it suffices to check the path in which player $b \in \{0, 1\}$ removes all resources from $R_b \cap P_b$, except the resources played in the outcome. We now show that 3-RESOURCE-DISARM^{NE} is weakly NP-hard by a reduction from the SUBSET-SUM problem.

Lemma 2. *3-RESOURCE-DISARM^{NE} is weakly NP-hard. This remains true even if the input outcome o^*, if feasible, would uniquely maximize welfare.*

Note that the last condition implies that finding a welfare-maximizing equilibrium of the disarmament game (rather than being given a specific o^* to achieve) is weakly NP-hard.

Proof. A SUBSET-SUM problem instance is defined by $K = \langle L, \vec{w}, W \rangle$, where $L = [n]^2$ is a set of goods and each good is associated with a positive value w_i, indicating good i has weight w_i. The objective is to determine whether there exists a subset $L' \subseteq L$ with total weight exactly W. Let $sumW = \sum_{i=1}^n w_i$; W.L.O.G, $W < sumW$. Given a SUBSET-SUM instance, we construct a multi-resource game $G(K) = \langle R_0, R_1, v_0, v_1 \rangle$ as follows:

- $R_0 = [n] \cup \{*\}$, $R_1 = \{x, y\}$;
- $v_0(i) = w_i - \varepsilon$ and $v_1(i) = -w_i - \varepsilon$ for all $i \in [n]$; moreover, $v_0(*) = 0$ and $v_1(*) = -W$;
- $v_0(x) = -sumW - \gamma$, $v_1(x) = sumW$, $v_0(y) = -sumW + W$, and $v_1(y) = 0$.

where $\gamma, \varepsilon > 0$ are sufficiently small. The target outcome is $o^* = (\emptyset, \emptyset)$, i.e., nobody uses any resources. (This does not necessarily mean that they *discarded* all the resources.) The condition in the theorem about maximizing welfare is satisfied with outcome o^*, because every resource gives negative social welfare.

---

[2] Let $[n]$ be the set $\{1, \cdots, n\}$; as a special case, $[0] = \emptyset$.

If there exists a solution L^* to the SUBSET-SUM problem K, then the on-path history with $h_1 = [R_0 \setminus L^*, R_1, 1]$, $h_2 = [R_0 \setminus L^*, \{y\}, 0]$ and $h_3 = [\emptyset, \{y\}, 1]$ is induced by a Nash equilibrium strategy profile with outcome o^*. One can check this on-path history satisfies the requirement of Lemma 1 by computing the security level of every non-terminal node according to Proposition 2: at the root, player 0's security level is $-n\varepsilon - \gamma - sumW + W < 0$; at h_1, player 1's security level is $-(n - |L^*|)\varepsilon + sumW - W - \sum_{i \in L \setminus L^*} w_i \leq -W + \sum_{i \in L^*} w_i = 0$; at h_2, player 0's security level is $-sumW + W + \sum_{i \in L \setminus L^*} w_i - \varepsilon < W - \sum_{i \in L^*} w_i = 0$; and at h_3 it is an equilibrium to play $(\emptyset, \emptyset)$.

For the other direction, suppose there exists an on-path history leading to o^* induced by a Nash equilibrium strategy profile. Note that $(\emptyset, \emptyset)$ cannot be a Nash equilibrium in the induced game if for the terminal node $z = [T_0, T_1]$, $L \cap T_0 \neq \emptyset$ or $x \in T_1$. Suppose the state after player 0's first move is $[R_0 \setminus R_0', R_1, 1]$. We first argue that $* \notin R_0'$. For the sake of contradiction, suppose $* \in R_0'$; then it must be that $R_0' = \{*\}$, because otherwise $sec_1(G_{R_0 \setminus R_0', R_1}) > sumW - \sum_{i \in [n] \setminus R_0'} (w_i + \gamma) > 0$. But, letting R_1' denote what 1 discards next, it cannot be the case that $x \in R_1'$, because this would result in $sec_0(G_{R_0 \setminus R_0', R_1 \setminus R_1'}) \geq \sum_{i \in [n]} w_i - sumW + W - n\varepsilon > 0$. But we have already shown that player 1 needs to discard x at this point. Hence we have our desired contradiction and can conclude that $R_0' \subseteq [n]$.

Again using Lemma 1, we must have $sec_1(G_{R_0 \setminus R_0', R_1}) = -W - (n - |R_0'|)\varepsilon + \sum_{i \in R_0'} w_i \leq 0$, that is, $\sum_{i \in R_0'} w_i \leq W + (n - |R_0'|)\varepsilon$. As for player 1's action in state $[R_0 \setminus R_0', R_1, 1]$, notice that player 1 cannot remove resources x and y together; otherwise, the security level of player 0 will be larger than 0. Thus, player 1 must remove only resource x at this turn, leading to state $[R_0 \setminus R_0', \{y\}, 0]$. Again using Lemma 1, we must have $sec_0(G_{R_0 \setminus R_0', \{y\}}) = -\sum_{i \in R_0'} w_i + W - (n - |R_0'|)\varepsilon \leq 0$, that is, $\sum_{i \in R_0'} w_i \geq W - (n - |R_0'|)\varepsilon$. With sufficiently small ε, these two inequalities turn to $\sum_{i \in R_0'} w_i = W$. Therefore, R_0' is a solution to the SUBSET-SUM instance. $\square$

Although K-RESOURCE-DISARM^{NE} is weakly NP-hard, it can be solved using dynamic programming to obtain a pseudopolynomial algorithm.

Theorem 2. *K-RESOURCE-DISARM^{NE} is weakly NP-complete. This remains true even if the input outcome o^*, if feasible, would uniquely maximize welfare.*

When there exists no limitation on the number of rounds, the problem becomes strongly NP-complete.

Theorem 3. *RESOURCE-DISARM^{NE} is strongly NP-complete. This remains true even if the input outcome o^*, if feasible, would uniquely maximize welfare.*

Proof. A 3-PARTITION problem instance is defined by a set of integers $I = \{w_1, \cdots, w_q\}$ with q divisible by 3; the objective is to partition $\{1, \cdots, q\}$ into subsets $W_1, \cdots, W_{q/3}$ such that for each i, $\sum_{j \in W_i} w_j = B$ where

$B = (3/q) \sum_{j=1}^q w_j$. 3-PARTITION remains strongly NP-complete even when we have $B/4 < w_j < B/2$ for each j, in which case we must have $|W_i| = 3$ for each i. We reduce from this special case, also assuming W.L.O.G that $B > 0$ (and hence all the w_j are positive too). Construct a multi-resource game $G(I) = \langle R_0, R_1, v_0, v_1 \rangle$ as follows:

- $R_0 = [q] \cup \{x, y\}$, $R_1 = \{z_1, z_2, \cdots, z_{q/3}\} \cup \{*\}$;
- $v_0(i) = w_i$ and $v_1(i) = -w_i - \gamma$ for all $i \in [q]$; moreover, $v_0(x) = 0$, $v_1(x) = -B$, $v_0(y) = \varepsilon - \delta$, and $v_1(y) = -\delta$;
- $v_0(z_j) = -B - 3\varepsilon/q$ and $v_1(z_j) = B$ for all $j \in [q/3]$; moreover, $v_0(*) = -\varepsilon$ and $v_1(*) = 0$;

where $\varepsilon/2 < \delta < \varepsilon$, $0 < \gamma, \varepsilon \ll B$, and the target outcome is $o^* = (\emptyset, \emptyset)$.

The condition in the theorem about maximizing welfare is satisfied with outcome o^*, again because every resource gives negative social welfare.

If there exists a solution $W_1^*, \cdots, W_{q/3}^*$ to the 3-PARTITION instance I, then consider the on-path history with $h_{2k-1} = [R_0 \setminus \bigcup_{j=1}^k W_j^*, R_1 \setminus \{z_1, \cdots, z_{k-1}\}, 1]$, $h_{2k} = [R_0 \setminus \bigcup_{j=1}^k W_j^*, R_1 \setminus \{z_1, \cdots, z_k\}, 0]$ for $1 \leq k \leq q$ and $h_{2q+1} = [\{x\}, \{*\}, 1]$. We argue that this corresponds to a a Nash equilibrium strategy profile with outcome o^*, by checking that this on-path history satisfies the requirement of Lemma 1 by computing the security level at every non-terminal node and appealing to Proposition 2. For $1 \leq k \leq q$, at h_{2k-1}, player 1's security level is $-(q/3 - k+1)(B+3\gamma) - \delta + (q/3 - k+1)B < -\delta$; at h_{2k}, player 0's security level is $(q/3-k)B - (q/3-k)(B+3\varepsilon/q) - \delta = -\varepsilon - \delta + 3k\varepsilon/q < 0$; and at h_{2q+1}, it is an equilibrium to play $(\emptyset, \emptyset)$.

For the other direction, suppose there exists an on-path history leading to o^* induced by a Nash equilibrium strategy profile. Note that $(\emptyset, \emptyset)$ cannot be a Nash equilibrium in the induced game if for the terminal node $z = [T_0, T_1]$, $([q] \cup \{y\}) \cap T_0 \neq \emptyset$ or $\{z_1, z_2, \cdots, z_{q/3}\} \cap T_1 \neq \emptyset$. Moreover, it is without loss of generality to assume that player 1 removes strategy $*$ in his final disarmament round if he does, since removing strategy $*$ earlier only increases player 0's security level for certain non-terminal nodes. Similarly, it is W.L.O.G to assume that player 0 removes strategy x in his final disarmament round if he does.

Consider the first iteration, i.e., the first disarmament step by both players. Suppose player 0 removes a subset R_0', and player 1 removes $c \geq 1$ strategies from $\{z_1, \cdots, z_{q/3}\}$ (note that both players have identical valuations for any of these). After player 0's move, by Lemma 1, the security level of player 1 must satisfy $-B + (q/3) * B - \delta - \sum_{i \in R_0 \setminus R_0'} (w_i + \gamma) = -B - \delta - (n - |R_0'|)\gamma + \sum_{i \in R_0'} w_i \leq 0$, that is, $\sum_{i \in R_0'} w_i \leq B + \delta + (n - |R_0'|)\gamma$. Moreover, after player 1's move, the security level of player 0 must satisfy $(q/3 - c)(-B - 3\varepsilon/q) - \delta + \sum_{i \in R_0 \setminus R_0'} w_i \leq 0$, that is, $\sum_{i \in R_0'} w_i \geq c(B + 3\varepsilon/q) - \varepsilon - \delta$. Combining the two inequalities (and the fact that $c \geq 1$), by the choice of ε, δ and the fact that $B/4 < w_i < B/2$ for all $i \in [q]$, we must

have $c = 1$ and $\sum_{i \in R'_0} w_i = B$ exactly. Repeating this argument inductively, we can show that for each iteration j, the subset R_0^j of strategies removed by player 0 must satisfy $\sum_{i \in R_0^j} w_i = B$ and player 1 removes one strategy from $\{z_1, \cdots, z_{q/3}\}$. Thus, the subsets $R_0^1, \cdots, R_0^{q/3}$ constitute a solution for the 3-PARTITION problem. $\square$

RESOURCE-DISARM$^{\text{SPNE}}$

We now turn to the computation of SPNE. We first obtain a version of Lemma 1 for the SPNE setting. Under SPNE, the players can only play an NE of the remaining multi-resource game even for the nodes not in the on-path history. By Proposition 1, in an NE of a multi-resource game (without disarmament), a player cannot use a resource that gives himself negative utility. Hence, such a resource has no impact at all under SPNE; in particular, it does not affect the other player's security level. (In contrast, in a non-subgame-perfect NE of the disarmament game, the resource may be used in parts of the disarmament game tree that are off the path of play, constituting a non-credible threat, and thereby help to implement a particular outcome.) Also, in an NE of a multi-resource game, a player must use any resource that provides himself positive utility. Hence, if such a resource also gives the other player positive utility, while in an NE of the disarmament game it is, off the path of play, possible to avoid playing that resource in order to punish the other player, this is not possible in an SPNE. We use the following notation for the set of resources that $-b$ will use to punish b in SPNE:

$$Q_{-b}(T_{-b}) = (P_{-b} \cap T_{-b}) \cup (N_b \cap (T_{-b} \setminus N_{-b})).$$

This is the union of the resources that $-b$ must use to best-respond in SPNE, and those that it is able to use in order to punish b. Define the *security level against a self-preserving player (SLAP)* as the maximum utility a player can guarantee himself against a player that is not willing to use resources that hurt himself.

Proposition 3. *The SLAP for player b in a two-player multi-resource game $G = \langle T_0, T_1, v_0, v_1 \rangle$ is*

$$SLAP_b(G_{T_0,T_1}) = \sum_{s \in P_b \cap T_b} v_b(s) + \sum_{s \in Q_{-b}(T_{-b})} v_b(s)$$

Now, we are ready to prove a similar lemma as Lemma 1 for the $SLAP$ level. Using this, we can characterize the computational complexity of 3-RESOURCE-DISARM^{SPNE} and RESOURCE-DISARM^{SPNE}.

Lemma 3. *P is an on-path history induced by a subgame perfect Nash equilibrium strategy profile with outcome o if and only if o is an equilibrium of the terminal node in P and for each non-terminal node $[T_0, T_1, b] \in P$, $SLAP_b(G_{T_0,T_1}) \leq u_b(o)$.*

Proof. "$\Rightarrow$": If o is not an equilibrium of the terminal node in P then we clearly do not have an SPNE. If there exists a non-terminal node $h = [T_0, T_1, b] \in P$ such that $SLAP_b(G_{T_0,T_1}) > u_b(o)$, then at node h, player b can deviate to choose Play and then play the strategy $P_b \cap T_b$ in

the induced game. By Proposition 1, player $-b$ must play a strategy X_{-b} for which $P_{-b} \cap T_{-b} \subseteq X_{-b} \subseteq T_{-b} \setminus N_{-b}$ in the induced game. Therefore, the utility of player b is $\sum_{s \in P_b \cap T_b} v_b(s) + \sum_{s \in X_{-b}} v_b(s) \geq \sum_{s \in P_b \cap T_b} v_b(s) + \sum_{s \in Q_{-b}(T_{-b})} v_b(s) = SLAP_b(G_{T_0,T_1}) > u_b(o)$.

"$\Leftarrow$": If o is an equilibrium of the terminal node in P and for each non-terminal node $h = [T_0, T_1, b] \in P$, $SLAP_b(G_{T_0,T_1}) \leq u_b(o)$, consider a strategy profile that specifies to:

- for nodes in the on-path history, choose the action to follow the on-path history; and for every non-terminal node not in the on-path history, choose Play;

- for each terminal node $z = [T_b, T_{-b}]$ not in P, player b plays $Q_b(T_b)$ and player $-b$ plays $Q_{-b}(T_{-b})$;

- for the terminal node $z = [T_b, T_{-b}]$ in P, both players play according to o.

First, note that for any subgame rooted at an off-path node, as in the proof of Theorem 1 (noting that $Q_b(T_b), Q_{-b}(T_{-b})$ form a Nash equilibrium of the multi-resource game $G_{T_b,T_{-b}}$), the strategy profile forms a Nash equilibrium. As for an on-path non-terminal node $[T_b, T_{-b}, b]$, if player b deviates to choose Play, according to the definition of the SLAP level, his utility is at most $SLAP_b(G_{T_b,T_{-b}})$ in the induced multi-resource game. If player b deviates to remove some resource, his utility can only decrease compared to choosing Play immediately. Finally, the players play the equilibrium o at the terminal node on the path. $\square$

Theorem 4. 3-RESOURCE-DISARM^{SPNE} *is weakly NP-complete and* RESOURCE-DISARM^{SPNE} *is strongly NP-complete. These remain true even if the input outcome o^*, if feasible, would uniquely maximize welfare.*

Fractional Disarmament

We now turn our attention to *fractional* disarmament, where resources are divisible and each player can reduce each resource by a fraction. In the previous section, the integral disarmament case, hardness results already appeared with two players. It turns out that with fractional disarmament, we can get certain positive results.

Example 2. *We now modify Example 1 so that each player only has resource x. Without resource y, it is impossible to successfully disarm in the integral disarmament model, because once a player removes x, the other player has a strict incentive to not remove his x.*

However, in the fractional disarmament model, successful disarmament is possible: at the k-th move of each player, he removes an additional fraction $1/2^k$ of resource x; in other words, he keeps a fraction $1/2^k$ of resource x after his k-th move. Assuming player 0 moves first, after player 0's k-th move, the equilibrium utility for ending disarmament at that point is $(-3/2^k, 0)$, while after player 1's k-th move, that utility is $(-1/2^k, -1/2^k)$. Therefore, both players are

incentivized to follow the path, and they can obtain utilities arbitrarily close to $(0,0)$ after sufficiently many steps.[3]

The above example demonstrates that with fractional disarmament, we can achieve objectives that cannot be obtained under integral disarmament. In fact, we show that any feasible outcome can be obtained for any number of players. Before formalizing our results, we first extend the definition of two-player multi-resource games to a multi-player version.

Definition 8. *An n-player multi-resource game is defined by $G = \langle R_1, \cdots, R_n, v_1, \cdots, v_n \rangle$, where player i's resource set is R_i and his valuation function is $v_i : \bigcup_{j=1}^{n} R_j \to \mathbb{R}$. Given strategy profile $(S_1, \cdots, S_n)$, the utility of player i is computed as $u_i(S_1, \cdots, S_n) = \sum_{s \in \bigcup_{j=1}^{n} S_j} v_i(s)$.*

For simplicity, assume $R_i = [m]$ for all $i \in [n]$ and let $\mathbf{0}$ (and $\mathbf{1}$) denote a vector of all zeroes (ones).

Fractional disarmament game We represent a remaining game in a terminal node as $\mathbf{p} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n]$. Here, $\mathbf{p}_i \in [0,1]^m$ represents the vector of the remaining fractions of player i's resources. We use $\mathbf{p}_i(s)$ to denote the s-th element in the vector $\mathbf{p}_i$. In the induced game $G_{\mathbf{p}}$, player i can play any vector $\mathbf{q}_i \in [0,1]^m$ with $\mathbf{q}_i \leq \mathbf{p}_i$. The utility function for player i becomes $u_i(\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n) = \sum_{j=1}^{n} \sum_{s \in R_j} \mathbf{q}_j(s) v_i(s)$.

To define a multi-player disarmament game $G_D^F(G)$, we introduce another option for the players during the disarmament stage (in addition to the options of Play and removing resources), namely a Pass action, which keeps the disarmament stage alive unless all players use it in sequence. The reason we need this action once we move beyond two players is that sometimes, in order to make progress, two players alternatingly need to remove actions while a third player waits and does nothing.

Therefore, we associate tuples $\{[\mathbf{p}_1, \cdots, \mathbf{p}_n, i, c] \mid \forall i \in [n], \mathbf{p}_i \in [0,1]^m; 0 \leq c < n\}$ with the non-terminal nodes, where i denotes the player to move and c denotes a counter that records the number of consecutive players that have chosen Pass. The counter is initialized to 0 and when a player chooses Pass, if the counter is less than $n-1$, the counter is incremented by 1 and the game turns to player $i+1$'s move; if it is equal to $n-1$, since all n players have chosen Pass, the game stage will be triggered by choosing Pass. If a player chooses to remove some fractions of resources, the counter is reset to 0. As in the previous (integral) definition of disarmament games, player i can still choose Play to trigger the game stage immediately, or choose a new vector of remaining resource fractions from $\{\mathbf{p}_i' \mid \mathbf{p}_i' \leq \mathbf{p}_i, \mathbf{p}_i' \in [0,1]^m\}$. The root of the game is $[\mathbf{1}, \cdots, \mathbf{1}, 1, 0]$.

Strategies, on-path histories, and outcomes in fractional disarmament games are defined similarly as before. Again,

let $P_i = \{s \in \bigcup_{j=1}^{n} R_j \mid v_i(s) > 0\}$ be the set of resources that give player i positive utility and $N_j = \{s \in \bigcup_{j=1}^{n} R_j \mid v_i(s) < 0\}$ the set of resources that give player i negative utility.

Security and SLAP level The notions of security level and SLAP are modified accordingly. Formally, the security level of player b in $G_{\mathbf{p}}$ is defined by

$$sec_i(G_{\mathbf{p}}) = \sum_{s \in P_i} \mathbf{p}_i(s) v_i(s) + \sum_{j \neq i} \sum_{s \in N_i} \mathbf{p}_j(s) v_i(s)$$

We use $Q_{i,j} = (R_j \cap P_j) \cup (N_i \cap (R_j \setminus N_j))$ to represent the set of resources that j will use to punish i in an SPNE. This is the union of the resources that j must use to best-respond in an SPNE, and those that it is able to use in order to punish i. Then we have

$$SLAP_i(G_{\mathbf{p}}) = \sum_{s \in P_i} \mathbf{p}_i(s) v_i(s) + \sum_{j \neq i} \sum_{s \in Q_{i,j}} \mathbf{p}_j(s) v_i(s)$$

In fractional disarmament games, Lemmas 1 and 3 still hold.

A folk theorem

Deng and Conitzer (2017) shows a folk theorem that all desirable utility profiles can be obtained for two-player disarmament games in which individual *mixed* strategies can be removed.[4] In that result, desirable utility profiles consist of utility profiles that are *feasible*—i.e., the utility profile results from some mixed-strategy profile—and *enforceable*—i.e., for each player, his utility is higher than his security level at the root of the disarmament game. However, it is not immediately clear how to extend this result to general games with more than two players and/or SPNE. For fractional disarmament games based on multi-resource games, however, we obtain a folk theorem result for games with more than two players and obtain a better convergence rate to approximate the equilibrium (whether NE or SPNE is the solution concept).

In fractional disarmament games, a utility profile $(u_1^*, \cdots, u_n^*)$ is feasible if and only if there exists a strategy profile $(\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*)$ such that for all $i \in [n]$: $u_i(\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*) = u_i^*$ and $\forall j \in [m], \mathbf{p}_i^*(j) > 0$ only if $v_i(j) \geq 0$.[5]

Theorem 5. *In fractional disarmament games $G_D^F(G)$, for any vector of feasible utilities $(u_1^*, \cdots, u_n^*)$ and its associated strategy profile $(\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*)$ with $u_i^* > sec_i(G_{\mathbf{1},\cdots,\mathbf{1}})$ for all $i \in [n]$, there exists a Nash equilibrium of $G_D^F(G)$ such that the terminal node of its induced on-path history is $z = [\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*]$ and the outcome o obtained at the terminal node satisfies $u_i(o) = u_i^*$ for all $i \in [n]$.*

---

[3]Of course, if it is common knowledge that a certain round k is the *last* disarmament round, then the last player to move has a strict incentive not to disarm at that point. This problem can be circumvented by tossing a public coin in each round that with small probability will signal the end of all disarmament, so that players are never sure that their current step is the last. This is discussed in more detail by Deng and Conitzer (2017).

[4]The removal of mixed strategies is technically similar to the fractional removal we consider here, though in our context it is more natural to think about resources being divisible rather than removing probability distributions.

[5]The folk theorem result in (Deng and Conitzer 2017) does not require the last condition. It does not need to, because that folk theorem requires that it is possible to remove fractions of *any* strategy. In our context, however, it is not possible to, say, remove the strategy of playing no resources at all. This makes it impossible to get someone to play a resource that harms himself.

Theorem 6. *In fractional disarmament games $G_D^F(G)$, for any vector of feasible utilities $(u_1^*, \cdots, u_n^*)$ with $u_i^* > sec_i(G_{\mathbf{1}, \cdots, \mathbf{1}})$ for all $i \in [n]$, there exists a $V\varepsilon$ (additive)-Nash equilibrium of $G_D^F(G)$, where V is the sum of absolute values associated with the resources, such that the length of its on-path history is $O(n^2 \log \frac{1}{\varepsilon})$ and the outcome o obtained at the terminal node satisfies satisfies $u_i(o) > u_i^* - V\varepsilon$ for all $i \in [n]$.*

All the results appear in this section also hold for SPNE if we replace the security level with the SLAP level.

A constructive proof

In this subsection, we provide a constructive proof for Theorems 5 and 6 via an algorithm to generate the corresponding on-path history. We first demonstrate an essential property of fractional disarmament games.

Lemma 4. *Given a multi-resource game and utility profile $(u_1^*, \cdots, u_n^*)$ for every player i, let $F = \{[\mathbf{p}_1, \cdots, \mathbf{p}_n] \mid \forall i \in [n], \mathbf{p}_i \in [0,1]^m, sec_i(G_{\mathbf{p}_1, \cdots, \mathbf{p}_n}) \leq u_i^*\}$. Then, F is a convex set.*

Given target utility profile $(u_1^*, \cdots, u_n^*)$ with corresponding strategy profile $(\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*)$, let $\mathbf{p}_i' = \mathbf{1} - \mathbf{p}_i^*$. We introduce parameter θ_i for each player i, initialized to 1, such that the vector of fractions of player i's remaining resources is $\mathbf{p}_i = \mathbf{p}_i^* + \theta_i \mathbf{p}_i'$. For convenience, let $G_{\theta_1, \cdots, \theta_n}^\theta = G_{\mathbf{p}_1^* + \theta_1 \mathbf{p}_1', \cdots, \mathbf{p}_n^* + \theta_n \mathbf{p}_n'}$. Define an update function f:

$$f_i(\theta_{-i}) = \inf\{\theta_i \geq 0 \mid \forall j, sec_j(G_{\theta_i, \theta_{-i}}^\theta) \leq u_j^*\}$$

That is, given θ_{-i}, function f returns the minimum θ_i such that every player's security level in the induced game is still lower than or equal to his target utility. We now present the algorithm to generate the on-path histories for Theorem 5 and 6. The idea behind the algorithm is to, in each stage, decrease all players' θ_i to the same value l, one by one, until we achieve the desired ϵ. This avoids getting stuck at a boundary but requires us to be conservative in choosing l.

Algorithm 1: Generate on-path strategy for feasible utilities that exceed security levels

Input: A multi-resource game G, a target utility profile $(u_1^*, \cdots, u_n^*)$ associated with its strategy profile $(\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*)$, and the approximation ratio ε

Output: An on-path history P for $G_D^F(G)$

For all $i \in [n]$, let $\mathbf{p}_i' = \mathbf{1} - \mathbf{p}_i^*$, $\theta_i = 1$;
Let $h_0 = [\mathbf{1}, \cdots, \mathbf{1}, 1, 0]$, $t = 0$, $l = 1$;
while $l > \varepsilon$ **do**
    $l = ((n-1)l + \max_{i \in [n]} f_i(\theta_{-i}))/n$;
    for $i : 1$ *to* n **do**
        $\theta_i = l$; $t = t + 1$;
        $h_t = [\mathbf{p}_1^* + \theta_1 \mathbf{p}_1', \cdots, \mathbf{p}_n^* + \theta_n \mathbf{p}_n', i+1, 0]$;
$h_{t+1} = [\mathbf{p}_1^* + \theta_1 \mathbf{p}_1', \cdots, \mathbf{p}_n^* + \theta_n \mathbf{p}_n']$;
return h;

Lemma 5. *When setting $\varepsilon = 0$, Algorithm 1 produces an on-path history that is part of a Nash equilibrium and in the limit leads to the terminal node $z = [\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*]$.*

Lemma 6. *Given an ε, Algorithm 1 produces an on-path history with length $O(n^2 \log \frac{1}{\varepsilon})$ that is part of a Nash equilibrium and leads to a terminal node $z = [\mathbf{p}_1^* + \varepsilon' \mathbf{p}_1', \cdots, \mathbf{p}_n^* + \varepsilon' \mathbf{p}_1']$ with $\varepsilon' < \varepsilon$, in which $u_i(G_{\varepsilon', \cdots, \varepsilon'}^\theta) > u_i^* - V\varepsilon'$ for all $i \in [n]$.*

Computation of minimum rounds

In the previous subsection, we showed that for any desirable utility profile, there exists a Nash equilibrium of a fractional disarmament game achieving those utilities; moreover, to get an ε-approximate Nash equilibrium, we need at most $O(n^2 \log \frac{1}{\varepsilon})$ rounds. Still, we may wish to compute the minimum number of rounds. In fractional disarmament games, define the number of removal rounds of an on-path history $P = (h_0 = \text{root}, h_1, \cdots, h_K, h_{K+1} = z)$ to be[6]

$$K - |\{h_i \in P \mid \chi(h_i) = \text{Pass}\}|$$

Definition 9 (MIN-FRACTIONAL-DISARM$^\mathcal{C}$ problem)**.** *In MIN-FRACTIONAL-DISARM$^\mathcal{C}$, given a fractional disarmament game $G_D(G)^F$, an outcome $o^* = (\mathbf{p}_1^*, \cdots, \mathbf{p}_n^*)$, and a parameter K, the objective is to determine whether there exists an equilibrium $(\sigma_1^*, \cdots, \sigma_n^*)$ of type $\mathcal{C}$ such that the outcome is o^* and the number of removal rounds in its induced on-path history is at most K.*

Unlike in the case where resources must be removed integrally, the problem is easy with two players by implementing the following scheme: (1) For a resource of player $b \in \{0,1\}$ that gives non-negative utilities to both players, remove the required fraction in player b's first disarmament round; (2) For a resource of player $b \in \{0,1\}$ that gives non-positive utility to player b, remove the required fraction in player b's last disarmament round; (3) For the remaining resources that give player b positive utility and player $-b$ negative utility; remove the resources i with larger $-v_b(i)/v_{-b}(i)$ first until the security level of player $-b$ equals $u_{-b}(o^*)$ in each round.

Theorem 7. *In 2-player fractional disarmament games, MIN-FRACTIONAL-DISARM^{NE} is in P.*

However, when we have multiple players, the problem becomes NP-complete.

Theorem 8. *In n-player fractional disarmament games, MIN-FRACTIONAL-DISARM^{NE} is strongly NP-complete even if each player has only one resource.*

Future Research

Future research could investigate game representations other than the multi-resource games defined here, for example ones that allow interaction between resources. (Of course, any representation scheme that can compactly represent all multi-resource games will inherit the hardness results given here.) It could also attempt to make the model more realistic by modeling imperfect observation of others' disarmament

---

[6]We do not count the Pass rounds because we can immediately jump to the next player instead.

actions, the possibility of re-arming (potentially at a cost), etc. These steps all may contribute to these techniques becoming practical enough for real applications.

Acknowledgement

References

Chen, X.; Deng, X.; and Teng, S.-H. 2009. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM* 56(3).

Conitzer, V., and Sandholm, T. 2006. Computing the optimal strategy to commit to. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, 82–90.

Conitzer, V., and Sandholm, T. 2008. New complexity results about Nash equilibria. *Games and Economic Behavior* 63(2):621–641.

Daskalakis, C.; Goldberg, P.; and Papadimitriou, C. H. 2009. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing* 39(1):195–259.

Deng, Y., and Conitzer, V. 2017. Disarmament games. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*.

Gilboa, I., and Zemel, E. 1989. Nash and correlated equilibria: Some complexity considerations. *Games and Economic Behavior* 1:80–93.

Porter, R.; Nudelman, E.; and Shoham, Y. 2008. Simple search methods for finding a Nash equilibrium. *Games and Economic Behavior* 63(2):642–662.

Tambe, M. 2011. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press.

von Stengel, B., and Zamir, S. 2010. Leadership games with convex strategy sets. *Games and Economic Behavior* 69:446–457.