Algorithms for Finding Approximate Formations in Games

Patrick R. Jordan
Yahoo! Labs
4401 Great America Parkway
Santa Clara, CA 95054
prjordan@yahoo-inc.com

Michael P. Wellman
University of Michigan
Computer Science & Engineering
Ann Arbor, MI 48109-2121
wellman@umich.edu

Abstract

Many computational problems in game theory, such as finding Nash equilibria, are algorithmically hard to solve. This limitation forces analysts to limit attention to restricted subsets of the entire strategy space. We develop algorithms to identify *rationally closed* subsets of the strategy space under given size constraints. First, we modify an existing family of algorithms for rational closure in two-player games to compute a related rational closure concept, called *formations*, for n-player games. We then extend these algorithms to apply in cases where the utility function is partially specified, or there is a bound on the size of the restricted profile space. Finally, we evaluate the performance of these algorithms on a class of random games.

Introduction

Researchers in multiagent systems often appeal to the framework of game theory, attracted by its well understood solution concepts and ability to model a wide variety of scenarios. However, many basic game-theoretic analysis operations, such as characterizing equilibria, are computationally complex (PPAD-complete or NP-complete). *Empirical game-theoretic analysis* (Wellman(2006)) inherits the complexity of algorithmic game theory and, in addition, concerns itself with the construction of game models from (costly) high-fidelity simulation. Given the dependence of computational cost on game size, reducing size—in particular by pruning strategies not relevant to analysis—can have a large impact on feasibility. However, determining if a strategy is irrelevant can itself be computationally complex, depending on the type of analysis undertaken.

Our goal is to find a *restricted* game that is minimal in size, yet conveys (approximately) the same relevant information as the base game. Obviously, what information is relevant depends on the exact type of strategic analysis, however one common vein in game-theoretic analysis is computing the *regret* of profiles. The regret computation is used to identify equilibria and other approximately stable profiles. In this context, we desire a restricted game such that the regret of any profile with respect to the restricted game is approximately the same as the regret of that profile with respect to the base game. This objective leads us to employ a *rational closure* concept called *formation*, defined below.

It is important to acknowledge that when we prune strategies, we lose the ability to calculate the regret of profiles not within the restricted game. As analysts, we can determine an acceptable tradeoff between this loss and the set of operations made feasible to us by the reduction.

Notation

A **strategic game** $\Gamma = \langle N, (S_i), (u_i) \rangle$ consists of a finite set of players, N, of size n indexed by i; a non-empty set of strategies S_i for each player; and a utility function $u_i : \times_{j \in N} S_j \rightarrow \mathbb{R}$ for each player. We use the symbol $\Gamma_{S \downarrow X}$ to denote a **restricted game** with respect to the **base game** Γ, where each player i in $\Gamma_{S \downarrow X}$ is restricted to playing strategies in $X_i \subseteq S_i$.

Each profile s is associated with the set of neighboring profiles that can be reached through a unilateral deviation by a player. The **unilateral deviation set** for player i and profile $s \in S$ is $\mathcal{D}_i(s) = \{(\hat{s}_i, s_{-i}) : \hat{s}_i \in S_i\}$, and the corresponding set unspecified by player is

$$\mathcal{D}(s) = \underset{i \in N}{\quad} \mathcal{D}_i(s).$$

Let $\Delta(\cdot)$ represent the probability simplex over a set. A **mixed strategy** σ_i is a probability distribution over strategies in S_i, with $\sigma_i(s_i)$ denoting the probability player i will play strategy s_i. The **mixed strategy space** for player i is given by $\Delta_i = \Delta(S_i)$. Similarly, $\Delta^S = \times_{i \in N} \Delta_i$ is the **mixed profile space**.

For a given player i, the best-response correspondence for a given profile σ is the set of strategies which yield the maximum payoff, holding the other players' strategies constant. Formally, the player i **best-response correspondence** for opponent profile $\sigma_{-i} \in \Delta(S_{-i})$ is

$$\mathcal{B}_i(\sigma_{-i}) = \underset{\hat{\sigma}_i \in \Delta_i}{\arg\max}\, u_i(\hat{\sigma}_i\ ,\ \sigma_{-i})$$

and for $\Delta \subseteq \Delta(S_{-i})$ is $\mathcal{B}_i(\Delta) = \cup_{\sigma_{-i} \in \Delta} \mathcal{B}_i(\sigma_{-i})$. The **overall best-response correspondence** for profile $\sigma \in \Delta(S)$ is $\mathcal{B}(\sigma) = \times_{i \in N} \mathcal{B}_i(\sigma_{-i})$ and for $\Delta \subseteq \Delta(S)$ is $\mathcal{B}(\Delta) = \cup_{\sigma \in \Delta} \mathcal{B}(\sigma)$.

The symbols $\mathfrak{B}_i(\sigma_{-i})$, $\mathfrak{B}_i(\Delta)$, $\mathfrak{B}(\sigma)$, and $\mathfrak{B}(\Delta)$ correspond to the pure-strategy variants of the best-response correspondences—allowing for a slight abuse of notation:

$$\begin{aligned}\mathfrak{B}_i(\sigma_{-i}) &= \mathcal{B}_i(\sigma_{-i}) \cap S_i,\\ \mathfrak{B}_i(\Delta) &= \mathcal{B}_i(\Delta) \cap S_i,\\ \mathfrak{B}(\sigma) &= \times_{i\in N}\mathfrak{B}_i(\sigma_{-i}),\\ \mathfrak{B}(\Delta) &= \cup_{\sigma\in\Delta}\mathfrak{B}(\sigma).\end{aligned}$$

We introduce symbols for the pure-strategy best-response to a set of profiles $X = \times_{i\in N}X_i$ where $\emptyset \subset X_i \subseteq S_i$:

$$\begin{aligned}\mathfrak{B}_i(X_{-i}) &= \mathfrak{B}_i\left(\times_{j\in N\setminus\{i\}}\Delta(X_j)\right),\\ \mathfrak{B}_i^\dagger(X_{-i}) &= \mathfrak{B}_i\left(\Delta(X_{-i})\right),\\ \mathfrak{B}(X) &= \times_{i\in N}\mathfrak{B}_i(X_{-i}),\\ \mathfrak{B}^\dagger(X) &= \times_{i\in N}\mathfrak{B}_i^\dagger(X_{-i}).\end{aligned}$$

The best-response correspondences $\mathfrak{B}(\cdot)$ and $\mathfrak{B}^\dagger(\cdot)$ have differing independence assumptions on *opponent mixtures* (Bernheim(1984)). Under $\mathfrak{B}(\cdot)$, each player's strategies are best responses to independent mixtures over opponent strategies; whereas under $\mathfrak{B}^\dagger(\cdot)$, the strategies are best responses to joint (correlated) mixtures over opponent strategies.

The regret measures described in this section quantify the stability of strategies and profiles, respectively. **A player's regret**, $\epsilon_i(\sigma_i \mid \sigma_{-i})$, for playing strategy $\sigma_i \in \Delta_i$ against opponent profile $\sigma_{-i} \in \Delta(S_{-i})$ is

$$\epsilon_i(\sigma_i \mid \sigma_{-i}) = \max_{s_i\in S_i} u_i(s_i, \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}).$$

Note that this can equivalently be formulated as $u_i(\hat{\sigma}_i, \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i})$ for any $\hat{\sigma}_i \in \mathcal{B}_i(\sigma_{-i})$. Given an efficient method for calculating a best-response, the latter formulation can speed up regret computations, especially if player i's strategy set is large or infinite.

Finally, we use the regret of the constituent strategies to define the regret of a profile. The **regret of profile** $\sigma \in \Delta$, is the maximum gain from deviation from σ by any player. Formally, $\epsilon(\sigma) = \max_{i\in N}\epsilon_i(\sigma_i \mid \sigma_{-i})$. A **Nash equilibrium** (NE) is a profile $\sigma \in \Delta^S$ such that $\sigma \in \mathcal{B}(\sigma)$, therefore $\epsilon(\sigma) = 0$. (Nash(1951)) proved that such an equilibrium always exists for finite games.

Pruning Strategies

From a player's perspective, the regret of playing a particular strategy in a profile depends on the other available strategies in the game. For a given profile, only the strategies that are improving deviations give the player positive regret and thus are relevant to the regret calculation. Suppose we can identify a joint strategy set such that, for each player and each available profile in the restricted space, the player's improving deviations are contained in the restricted set of strategies for that player. Strategies outside that set are irrelevant from the point of the regret analysis. The restricted game formed from the joint strategy sets is *closed*, in the sense that a rational player would not choose to play strategies outside those of the restricted game, given knowledge that the other players choose strategies within their restricted strategy sets. On that basis, we can prune these irrelevant strategies and reduce the size of the game.

Eliminating strategies can be used as a pre-processing step to reduce the complexity of many game-theoretic algorithms, such as finding equilibria. For instance, *iterated elimination of strictly dominated strategies* (IESDS) removes dominated strategies from the players' strategy sets. Dominated strategies are *non-rationalizable* (Bernheim(1984); Pearce(1984)), thus not contained in any equilibrium. Another possibility is to compute all of the Nash equilibria of the game and keep only the strategies participating in some equilibrium. However, if identifying equilibria is the goal, then computing them as a preprocessing step does not reduce the complexity of the analysis as a whole.

(Conitzer and Sandholm(2005)) introduce a general *eliminability criterion* for two-player games. The authors compute whether a strategy is eliminable by solving a mixed integer program, which implicitly considers the rationalizable and NE solution concepts discussed so far. Whereas iterated elimination using this criterion cannot rule out equilibrium strategies in the base game, it may introduce NE in the restricted game that are not equilibria in the base game. This renders it difficult to identify NE of the base game directly from equilibrium analysis of the restricted game.

Rational Closure

A restricted strategy set, X, is *rationally closed* if the set is closed under a given best-response correspondence. In this section, we review two formal definitions of rational closure from prior literature.

Definition 1 ((Basu and Weibull(1991))). *A set of profiles $X \subseteq S$ is*

- ***closed under rational behavior** (CURB) if $\mathfrak{B}(X) \subseteq X$.*
- *a **minimal CURB set** if no proper subset of X is CURB.*

Definition 2 ((Harsanyi and Selten(1988))). *A set of profiles $X \subseteq S$ is*

- *a **formation** if $\mathfrak{B}^\dagger(X) \subseteq X$.*
- *a **minimal formation**[1] if no proper subset of X is a formation.*

The subtle distinction between these two concepts lies in the use of different best-response correspondences: $\mathfrak{B}^\dagger(X)$ comprises best responses to correlated opponent mixtures, whereas strategies in $\mathfrak{B}(X)$ are best responses to independent mixtures.

Each formation contains at least one Nash equilibrium in the base game. Moreover, the regret of a profile in a formation, with respect to the restricted game defined by the formation, is equal to the regret of the profile with respect to the base game.

Fact 1. *If X is a formation and $\sigma \in \Delta^X$, then $\epsilon(\sigma \mid \Gamma) = \epsilon(\sigma \mid \Gamma_{S\downarrow X})$.*

[1](Harsanyi and Selten(1988)) call this a *primitive* formation, but we use *minimal* to emphasize the parallel with minimal CURB sets.

Proof. By construction, for each player $i \in N$, $S_i \supseteq X_i \supseteq \mathfrak{B}_i(\sigma_{-i})$. Therefore,

$$\begin{aligned}\epsilon(\sigma \mid \Gamma) &= \max_{i\in N}\max_{s_i\in S_i} u_i(s_i,\sigma_{-i}) - u_i(\sigma)\\ &= \max_{i\in N}\max_{s_i\in \mathfrak{B}_i(\sigma_{-i})} u_i(s_i,\sigma_{-i}) - u_i(\sigma)\\ &= \max_{i\in N}\max_{s_i\in X_i} u_i(s_i,\sigma_{-i}) - u_i(\sigma)\\ &= \epsilon(\sigma \mid \Gamma_{S\downarrow X}).\end{aligned}$$

□

(Benisch et al.(2006)Benisch, Davis, and Sandholm) introduce a set of algorithms for efficiently finding CURB sets in two-player games. In the following section, we describe an extension of their algorithm for n-player games that identifies the minimal formation sets. Because $\mathfrak{B}(X) \subseteq \mathfrak{B}^{\dagger}(X)$, minimal formations weakly contain minimal CURB sets. Minimum formations are also closely related to saddle points (Shapley(1964)). In fact, the set of *strict* mixed saddles (Duggan and Le Breton(2001)) is equivalent to the set of minimal formations. (Brandt et al.(2009)Brandt, Brill, Fischer, and Harrenstein) show that the set of all strict mixed saddles can be computed in time polynomial in $|S|$.

Finding Minimal Formations

In this section we describe a minimum-formation-finding algorithm, called MCF, which is an n-player extension of Min-Containing-CURB by (Benisch et al.(2006)Benisch, Davis, and Sandholm). Given the equivalence of strict mixed saddles and minimal formations, MCF turns out to be identical to the minMGSP algorithm of (Brandt et al.(2009)Brandt, Brill, Fischer, and Harrenstein). Rather than describe the algorithms on the basis of dominance like (Brandt et al.(2009)Brandt, Brill, Fischer, and Harrenstein), we proceed on the basis of rationalizability to ease exposition in the subsequent sections, where we address partially-specified games and approximate formations.

(Benisch et al.(2006)Benisch, Davis, and Sandholm) give algorithms for finding all minimal CURB sets, a sample minimal CURB set, and the smallest minimal CURB set. All three algorithms rely on a subroutine that computes $\mathfrak{B}_i(X)$ for some $X \subseteq S$. The authors determine these strategies by solving a *feasibility* problem, in essence checking whether each strategy is a best response to some opponent mixture. However, when players mix over strategies independently, the feasibility problem is nonlinear for more than two players. Therefore, the authors restrict attention to two-player games.

We describe an extension of their algorithms that is instead based on the computation of $\mathfrak{B}_i^{\dagger}(X)$. Allowing correlated opponent play (replacing CURB sets with formations) restores the linearity of the feasibility problem. The Correlated-All-Rationalizable (CAR) algorithm extends the All-Rationalizable algorithm (Benisch et al.(2006)Benisch, Davis, and Sandholm) to n-player games with correlated opponent play; correspondingly, the Minimum-Containing-Formation (MCF) algorithm is the extension of the two-player Min-Containing-CURB algorithm. Like the two-player algorithm, MCF uses seed strategies to generate a minimum formation containing the seed strategy. In many cases, the minimum formation is also the minimum CURB set containing the seed strategies.

Algorithm 1 CAR(S_i, X_{-i}, u_i)

```
S_i* ← ∅
for s_i ∈ S_i do
    if solution to this linear feasibility program exists
    find σ_{-i} such that
                              Σ_{x_{-i} ∈ X_{-i}} σ_{-i}(x_{-i}) = 1
        (∀x_{-i} ∈ X_{-i})    σ_{-i}(x_{-i}) ≥ 0
        (∀ŝ_i ∈ S_i)          u_i(s_i, σ_{-i}) ≥ u_i(ŝ_i, σ_{-i})
    then
        S_i* ← S_i* ∪ {s_i}
return S_i*
```

Algorithm 2 MCF($s, \langle N, (S_i), (u_i)\rangle$)

```
for j ∈ N do
    S_j* ← {s_j}
converged ← false
while ¬converged do
    converged ← true
    for j ∈ N do
        Ŝ_j ← CAR(S_j, ×_{k∈N\{j}} S_k*, u_j)
        if Ŝ_j \ S_j* ≠ ∅ then
            S_j* ← S_j* ∪ Ŝ_j
            converged ← false
return ⟨S_1*, ..., S_n*, u⟩
```

Algorithms for finding all minimal formations, a sample minimal formation, and the smallest minimal formation are constructed by substituting the MCF algorithm for the Min-Containing-CURB algorithm in the respective CURB-set procedure (Benisch et al.(2006)Benisch, Davis, and Sandholm). Seed strategies must be provided for $n-1$ of the players in the non-symmetric cases and one player in the symmetric case. All algorithms have time complexity polynomial in $|S|$.

In MCF, we see the connection between minimal formations and strict mixed saddles. Where the MCF algorithm computes correlated rationalizable strategies, the minMGSP algorithm computes strictly undominated strategies. Under correlated opponent play, the set of rationalizable strategies is equivalent to the set of strictly undominated strategies; thus, the minMGSP algorithm is identical to MCF.

We investigate extensions of the minimal formation algorithms for settings that occur in empirical game-theoretic analysis. These extensions modify the linear feasibility program (LFP) of CAR. First, we extend the algorithms for the case when the utility function is partially specified. Second, we introduce algorithms to identify approximate formations when we have a bound on the size of the game that can be analyzed.

Partially Specified Games

For a fully specified game Γ, the utility $u(s)$ is defined for every profile $s \in S$. In some contexts, for example when inducing game models from empirical observations (Jordan and Wellman(2009)), we may have only a partial game specification, where utility is evaluated for a strict subset of profiles. In this section, we explore methods of assigning utility estimates to the *missing profiles* M, and the implications of these methods for the minimal formation algorithms of the previous section. We consider the domain of $u(\cdot)$ to be $S \backslash M$ and, if $M \neq \emptyset$, we call $u(\cdot)$ a *partially specified utility function* (PSU).

The CAR algorithm uses the payoffs given by the utility function $u(\cdot)$ to determine which strategies are best responses to mixtures over X_{-i}. The algorithm assumes that we have an estimate for the utility of each profile $s \in S_i \times X_{-i}$, however this may not be the case. First, consider the use of a *default utility* $\tilde{u}$ for $s \in M$.

If $s \in M$, we use $\tilde{u}$, in place of $u_i(s)$, when solving the feasibility problem in the CAR algorithm. This may give an unduly pessimistic view of the relative utilities associated with a strategy $s_i \in S_i$, when determining if s_i is a best response. This can occur in two ways for a given opponent profile $x_{-i} \in X_{-i}$. First, $(s_i, x_{-i}) \in M$ and the actual utility $u(s_i, x_{-i}) > \tilde{u}$. Second, for some $\hat{s}_i \in S_i \setminus \{s_i\}$, $(\hat{s}_i, x_{-i}) \in M$ and the actual utility $u(\hat{s}_i, x_{-i}) < \tilde{u}$. Given either of those scenarios, we could exclude s_i from S_i^* given our current partially specified utility function $u(\cdot)$, only to learn later that s_i is rationalizable for some mixture over X_{-i} once we have estimates for the associated utilities. Therefore, once again, we could have a profile that is an equilibrium in the restricted game, but not in the base.

We would like to eliminate as many strategies as our PSU sanctions, but not at the expense of removing strategies in the support of a minimal formation. One effect of erroneous strategy elimination is the loss of the ability to make general claims about NE using only the subset of the profile space returned by the formation-finding algorithms. Notice that we can straightforwardly check whether $X \subseteq S$ is a formation, if $\mathcal{D}(X) \cap M = \emptyset$. However, restricting formation analysis to subspaces with estimated utilities is too strong a requirement.

Instead of using a default utility, if we make *optimistic assumptions* on the utilities of missing profiles, we can still guarantee the sets returned by any of the formation finding algorithms are actually formations, albeit not necessarily minimal formations. As in the consideration of default utility values, we focus on the relative utilities of a strategy $s_i \in S_i$. Let $[u_i^-, u_i^+]$ be known bounds on the utility of player i. In the algorithm, the utilities for player i of the missing profiles are assigned to either u_i^- or u_i^+, according to the following two rules:

- if $(s_i, x_{-i}) \in M$, then u_i^+ is used for $u_i(s_i, x_{-i})$;
- for $\hat{s}_i \in S_i \setminus \{s_i\}$, if $(\hat{s}_i, x_{-i}) \in M$, then u_i^- is used for $u_i(\hat{s}_i, x_{-i})$.

Let MCF-PSU be the extension of the MCF algorithm using the previous rules for missing profiles.

Fact 2. *If X is the set returned by MCF and X_{PSU} is the set returned by MCF-PSU for the same parameters, then $X \subseteq X^{PSU}$.*

Proof. We use induction on the size of the missing profile set. For the **base case**, let $M = \{s\}$. Assume that $X \not\subseteq X^{\text{PSU}}$. Therefore for some player i, there is a strategy $s_i^* \in X_i$ that is not in X_i^{PSU}. Because $s_i^* \in X_i$, there is some iteration in MCF such that the LFP is satisfied for s_i^* in the CAR algorithm. Let p^* be the probability vector that solves the LFP. Three cases can occur: (i) s is a profile in the left-hand-side (LHS) of the utility-based constraints, (ii) s is a profile in the right-hand-size (RHS), (iii) s is neither a profile in the LHS nor the RHS. In case (i), $u_i(s) \leq u_i^+$ implies that the LHS constraints are weakly greater under CAR-PSU, and p^* satisfies the new LFP. In case (ii), $u_i(s) \geq u_i^-$ implies that the RHS constraints are weakly less under CAR-PSU, and p^* satisfies the new LFP. In case (iii) the CAR-PSU utilities are unchanged, so p^* trivially satisfies the LFP. Therefore, the PSU-based LPF is satisfied for s_i^* under CAR-PSU and $s_i^* \in X_i^{\text{PSU}}$ Hence, by contradiction, we have $X \subseteq X^{\text{PSU}}$.

For the **inductive step**, assume $X \subseteq X^{\text{PSU}}$ where M_k is the missing profile set and $|M_k| = |k|$. We use the same reasoning to conclude that under $M_k + \{s\}$, $X \subseteq X^{\text{PSU}}$: the LHS (RHS) constraints weakly increase (decrease) when under $M_k + \{s\}$. Therefore, by induction, $X \subseteq X^{\text{PSU}}$ under any countable M. □

Under these rules, s_i can never be eliminated with the *optimistic partially specified utility function*, when it would not have been with the fully specified utility function. In addition, s_i cannot support the elimination of another strategy with the optimistic partially specified utility function, when it would not have been supporting with the fully specified utility function. Therefore, each minimal formation is weakly contained within some formation found using the optimistic algorithm.

Approximate Formations

In empirical game-theoretic analysis, it may be feasible (due to the cost of simulation) only to analyze a game of bounded size. Given such a bound, we may endeavor to find a restricted game that includes near-equilibrium profiles of the base game, such that their regrets in the restricted and base games are approximately the same. In this section, we develop techniques for finding such strategy sets given a fully-specified utility function, however these techniques can be applied to scenarios where the utility function is partially specified.

For instance, (Jordan et al.(2010)Jordan, Schvartzman, and Wellman) describe the *strategy exploration problem*, in which modelers attempt to determine a simulation sequence for strategy sets. In this problem's formulation, *simulating* a set of strategies reveals the utilities of all supported profiles. Thus, in the terminology of the previous section, the utility function is partially specified, and at each step in the exploration sequence the profiles whose utilities are revealed are removed from the missing profile set. The modeler's goal in the strategy exploration problem is to minimize the regret (with respect to the base game) of the minimum-regret profile in the restricted strategy space formed after the terminal step. (Jordan et al.(2010)Jordan, Schvartzman, and Wellman) found that strategy exploration policies that heuristically minimize the regret of the explored strategy set (defined below) performed well. Algorithms that efficiently compute approximate formations (strategy sets with minimal regret) are critical to these policies. In this section, we describe our approach to computing these approximate formations given a bound on the size of the resulting strategy set.

Let $\widehat{U}_i$ be the function that specifies the best-response utility of player i for each $\sigma_{-i} \in \Delta(S_{-i})$ when player i's strategy set is limited to X_i. That is,

$$\widehat{U}_i(\sigma_{-i}; X_i) = \max_{s_i \in X_i} u_i(s_i, \sigma_{-i}).$$

We extend the definition of regret to sets of strategies in the following way. For $\emptyset \subset X_i \subseteq S_i$, the *regret* of player i for having the restricted strategy set X_i against $\Delta(X_{-i})$ is

$$\epsilon_i(X_i \mid X_{-i}) = \max_{\sigma_{-i} \in \Delta(X_{-i})} \widehat{U}_i(\sigma_{-i}; S_i) - \widehat{U}_i(\sigma_{-i}; X_i).$$

In other words, $\epsilon_i(X_i \mid X_{-i})$ represents a bound on the potential gain to player i for deviating from its restricted set X_i, in the context where other agents are restricted to X_{-i}. We similarly define regret of a joint strategy set. For $\emptyset \subset X \subseteq S$, the *regret* over all players for having restricted joint strategy set X is $\epsilon(X) = \max_{i \in N} \epsilon_i(X_i \mid X_{-i})$. A set of profiles $X \subseteq P$ is an ϵ-**formation** if $\epsilon(X) \leq \epsilon$. The set X is said to be an ϵ-**minimal formation** if no proper subset of X is an ϵ-formation.

We consider two general scenarios in which we minimize $\epsilon(X)$. In the first scenario, we are interested in finding a set of joint strategies with minimal regret (ϵ) under some budget k on the size of the profile space. Such a situation can occur when analysts are designing novel strategies from a set of existing strategies (Jordan et al.(2007)Jordan, Kiekintveld, and Wellman). In that work, new strategies are evaluated against a set of promising existing strategies using NE regret. To compute NE regret, modelers need estimates for the utility of all the profiles in the joint strategy space as well as the unilateral deviations to the new strategy. Modelers may be limited in the number of observations they can make for these new profiles, due to limited computational resources. This implicitly bounds the size of the joint-strategy space they consider. We formulate this optimization problem as follows:

$$\begin{array}{rl} \min & \epsilon(X) \\ \text{s.t.} & |X| \leq k. \end{array} \tag{1}$$

For the second scenario, consider the central heuristic in the formation-based strategy exploration policies given by (Jordan et al.(2010)Jordan, Schvartzman, and Wellman). At each step these policies determine a strategy to add into an existing strategy space. Each profile in the new profile space is simulated, causing the corresponding utilities to be revealed. The formation-based policies heuristically select a strategy that is a maximally beneficial deviation to a particular mixture over a minimum-regret formation in the explored strategy space. Let the explored strategy space be given by $E \subseteq S$. Thus, the first step in selecting the strategy involves finding a minimum-regret formation by solving

$$\begin{array}{rl} \min & \epsilon(X) \\ \text{s.t.} & X \subseteq E. \end{array} \tag{2}$$

In the remainder of the section, we develop heuristic algorithms for solving (1), noting that solving (2) involves only minor modifications to these algorithms.

In order to solve the optimization problem given by (1), we need to efficiently compute $\epsilon(\cdot)$, which in turn depends on the calculation of $\epsilon_i(\cdot)$. We calculate $\epsilon_i(X_i \mid X_{-i})$ by determining the supporting strategies in $S_i \setminus X_i$. A strategy $s_i \in S_i \setminus X_i$ *supports* ϵ if the maximum gain in utility $\tau > 0$ in the following linear program:

$$\begin{array}{rrcl} \max & \tau & & \\ \text{s.t.} & \sigma_{-i}(x_{-i}) & = & 1 \\ (\forall x_{-i} \in X_{-i}) & \sigma_{-i}(x_{-i}) & \geq & 0 \\ (\forall \hat{s}_i \in X_i) & u_i(s_i, \sigma_{-i}) - \tau & \geq & u_i(\hat{s}_i, \sigma_{-i}). \end{array} \tag{3}$$

Notice that the last set of constraints in the linear program of (3) is slightly different than the set of constraints in the linear feasibility program of Algorithm 1: $\hat{s}_i \in S_i$ in Algorithm 1, whereas $\hat{s}_i \in X_i$ in (3). This difference is due to the fact that player i is constrained to playing strategies in X_i.

Let COMPUTE-TAU return the solution to the linear program of (3). The pseudo-code for calculating $\epsilon_i(X_i \mid X_{-i})$ is given in Algorithm 3. For each $s_i \in S_i \setminus X_i$, we determine the maximum τ and set ϵ to the maximum of those values. If $\tau \leq 0$, then s_i is *covered* by X_i. If $\tau > 0$, then s_i *supports* $\epsilon_i(X_i \mid X_{-i})$.

Algorithm 3 COMPUTE-REGRET$(X_i, S_i, X_{-i}, u_i(\cdot))$

```
ε_i ← 0
for s_i ∈ S_i \ X_i do
    τ ← COMPUTE-TAU(X_i, s_i, X_{-i}, u_i(·))
    ε_i ← max(ε_i, τ)
return ε_i
```

Given an efficient algorithm for determining $\epsilon(\cdot)$, we need an efficient search algorithm for identifying the optimal restricted game. Before discussing search algorithms, consider some general observations regarding $\epsilon(\cdot)$: ϵ is not monotonic in the subset relation, ϵ is not transitive, ϵ is not submodular or supermodular, and greedy selection of X is not optimal. The first three observations provide no effective bounds for branch-and-bound search. The fourth observation rules out a greedy search algorithm as optimal, however it may be a reasonable heuristic search method. Below we outline a simple algorithm to search over the join-semilattice of restricted games.

Algorithm 4 uses two basic subroutines and a priority queue with a maximum size of Q to determine the optimal restricted game for the given bound k. The two enqueue subroutines, ENQUEUE-INITIAL-GAMES and ENQUEUE-CHILD-GAMES, each present restricted games to the priority queue. The ENQUEUE-INITIAL-GAMES subroutine generates all restricted games where each player has a single

Algorithm 4 FIND-FORMATION($\Gamma,\ k,\ Q$)

$best \leftarrow null$
$queue \leftarrow$ Empty priority queue with maximum size Q ordered by $\epsilon(\cdot)$
ENQUEUE-INITIAL-GAMES($queue, \Gamma,\ k$)
while *queue is not empty* **do**
    $game \leftarrow$ top($queue$)
    if *best is null or* $\epsilon(best) > \epsilon(game)$ **then**
        $best \leftarrow game$
    ENQUEUE-CHILD-GAMES($queue,\ game,\ k$)
return $best$

strategy to choose from. The ENQUEUE-CHILD-GAMES subroutine generates all of the restricted games where some player has an additional strategy to choose from. For each generated restricted game in the enqueue subroutines, the regret of the game is calculated by the COMPUTE-EPSILON subroutine. The priority queue orders each restricted game by this calculated value.

Observe that the number of restricted games is exponential in the size of the players' strategy sets, therefore a finite (small) value for Q is required for even small values of k and small games. Notice the special case of $Q = 1$, which corresponds to simple greedy search. In the experiments below, we vary Q and observe the effects on the regret of the selected restricted game.

Experiments

The randomly generated games used in this section are classified into two distinct types generated by GAMUT (Nudelman et al.(2004)Nudelman, Wortman, Shoham, and Leyton-Brown): *random* and *covariant*. The random class of games has payoffs that are uniformly and independently distributed in the range [-100,100]. The covariant class of games has payoffs that are distributed *Normal*[0,1] with covariance r between players in a profile. For experiments in this section, we used a setting of $r = -\frac{1}{2}$ for covariant games. We generated 100 instances of each class with two players and ten strategies per player. We compute the regret for each pure-strategy profile in each instance.

The approximate-formation algorithm seeks to find a restricted game whose regret is as small as possible for a given constraint k on the size of the game, measured in number of profiles. (Benisch et al.(2006)Benisch, Davis, and Sandholm) found experimentally that random games tend to have small smallest CURB sets (pure strategy equilibria), while covariant games have large smallest CURB sets (nearly all strategies). While reviewing their findings, we make another observation. For those random games which do not have a pure-strategy NE, the smallest CURB sets tend to be large. This implies that we are unlikely to find minimal CURB sets for intermediate values of k unless there exists a pure-strategy Nash equilibrium.

GAMUT generated games with PSNE around 58% of the time for random games and 14% of the time for covariant games. Therefore with the same frequency, we can find minimal CURB sets where each player is selecting a single strategy. Thus returning a restricted game consisting of the minimum-regret pure-strategy profile is optimal in nearly 58% and 14% of these games, respectively. Computing the minimum-regret profile is linear in the number of profiles and thus can be accomplished efficiently.

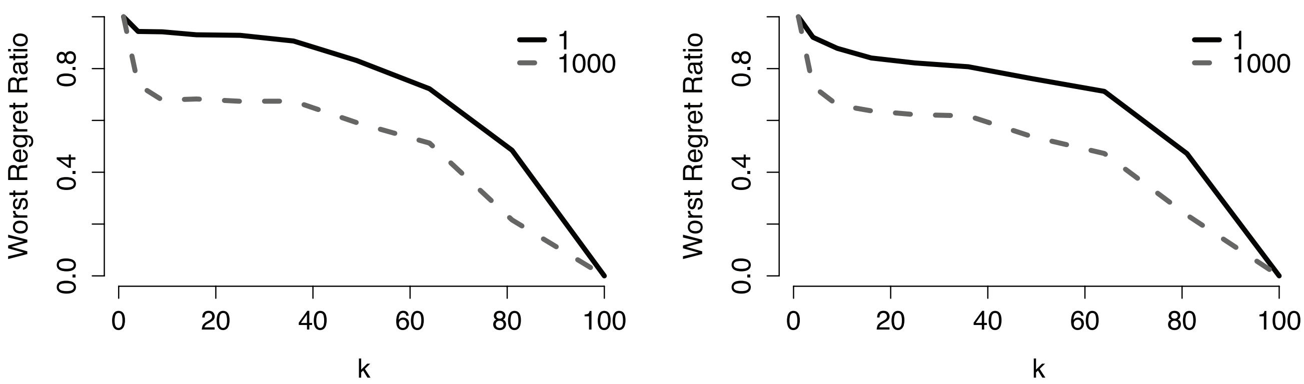
The question remains as to what should be done in the remaining 42% and 86% of respective cases when we are bounded by an intermediate value for k. From Benisch et al.'s results, we can conclude that it is unlikely that a minimum CURB set, an ϵ-formation with $\epsilon = 0$, is found when k is much smaller than the size of the base game. However, using our approximate formation algorithms, we can find ϵ-formations with $\epsilon > 0$ for any k.

The number of restricted games is exponential and we have no known bounds relating the regret of different restricted games. However, the Q parameter in Algorithm 4 allows us to explore some of the restricted games off of the greedy path if $Q > 1$. Therefore, we remove instances from the two classes where a PSNE exists. On the remaining games, we run Algorithm 4 with two settings for Q and various settings for k. For each game, we run the algorithm for each k, where $k \in \{i^2 \mid 1 \leq i \leq 10\}$. Because the size of the base game is 100, the algorithm should return a game with $\epsilon = 0$ when $k = 100$, even in the greedy case. We use two settings for Q; a setting of $Q = 1$ corresponds to greedy search, whereas $Q = 1000$ allows some non-greedy exploration. Maintaining a maximum queue size of 1000 allows a complete search of roughly the first two levels, or $k \leq 9$ in the generated instances.

Figure 1 shows the results of running Algorithm 4 on the two classes of games, where the *worst-case regret ratio* is fraction of regret of the restricted game found by the algorithm when compared to the regret of the minimum-regret profile. Both the greedy ($Q = 1$) and $Q = 1000$ found restricted games with less regret than the minimum regret profile. In fact, when considering the random and covariant game classes, we can find a restricted game with worst-case regret ratio of approximately 75% using only 9% of the base game space on average.

Conclusion

Formations provide a useful basis for selecting restricted games that capture relevant strategic information about the base game. In particular, formations allow modelers to be confident that the regret of any profile calculated in the restricted game will be no larger when calculated in the base game. We adapt the algorithms of (Benisch et al.(2006)Benisch, Davis, and Sandholm) for calculating minimal CURB sets in two-player games to finding minimal formations in n-player games (constructing a set of algorithms equivalent to the mixed-generalized-saddle-point algorithms of (Brandt et al.(2009)Brandt, Brill, Fischer, and Harrenstein)). This formation-based extension allows us to derive techniques for identifying minimal formations when utility functions are only partially specified, as in empirical game-theoretic analysis when the strategy space is large and sampling is costly. In addition, we provide a heuristic search



(a) Random (b) Covariant

Figure 1: Worst-case regret ratio for random and covariant games.

algorithm that determines an optimal ϵ-formation given either a bound on the size of the restricted strategy space or a constraint on the strategies in the space itself. The latter is of critical importance to the heuristic strategy exploration policies of (Jordan et al.(2010)Jordan, Schvartzman, and Wellman). In an experiment on random games, we show that the ϵ-formation finding algorithm is able to decrease regret compared to min-regret profiles using relatively small strategy spaces.

Acknowledgments

We are grateful to David Parkes and the anonymous reviewers for comments that helped us clarify the technical exposition and relation to prior work. This work was supported in part by the US National Science Foundation under grant CCF-0905139.

References

K. Basu and J. W. Weibull. Strategy subsets closed under rational behavior. *Economic Letters*, 36(2):141–146, 1991.

M. Benisch, G. B. Davis, and T. Sandholm. Algorithms for rationalizability and CURB sets. In *Twenty-First National Conference on Artificial Intelligence*, pages 598–604, Boston, 2006.

B. D. Bernheim. Rationalizable strategic behavior. *Econometrica*, 52(4):1007–1028, 1984.

F. Brandt, M. Brill, F. Fischer, and P. Harrenstein. Computational aspects of Shapley's saddles. In *Eighth International Conference on Autonomous Agents and Multi-Agent Systems*, pages 209–216, 2009.

V. Conitzer and T. Sandholm. A generalized strategy eliminability criterion and computational methods for applying it. In *Twentieth National Conference on Artificial Intelligence*, pages 483–488, Pittsburgh, 2005.

J. Duggan and M. Le Breton. Mixed refinements of Shapley's saddles and weak tournaments. *Social Choice and Welfare*, 18(1):65–78, 2001.

J. C. Harsanyi and R. Selten. *A General Theory of Equilibrium Selection in Games*. MIT Press, 1988.

P. R. Jordan, C. Kiekintveld, and M. P. Wellman. Empirical game-theoretic analysis of the TAC supply chain game. *Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1188–1195, 2007.

P. R. Jordan, L. J. Schvartzman, and M. P. Wellman. Strategy exploration in empirical games. In *Ninth International Conference on Autonomous Agents and Multiagent Systems*, Toronto, May 2010.

P. R. Jordan and M. P. Wellman. Generalization risk minimization in empirical game models. In *Eighth International Conference on Autonomous Agents and Multi-Agent Systems*, pages 553–560, Budapest, 2009.

J. Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295, 1951.

E. Nudelman, J. Wortman, Y. Shoham, and K. Leyton-Brown. Run the GAMUT: A comprehensive approach to evaluating game-theoretic algorithms. In *Third International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 880–887, New York, 2004.

D. G. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050, 1984.

L. S. Shapley. Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, volume 52 of *Annals of Mathematics Studies*, pages 1–29. Princeton University Press, 1964.

M. P. Wellman. Methods for empirical game-theoretic analysis (extended abstract). In *Twenty-First National Conference on Artificial Intelligence*, pages 1552–1555, Boston, 2006.