Abstraction and Approximation in Games: Analyzing the Exploitability Graph

Adrian Marple [a,1,*], Yoav Shoham[a,1]

[a]Stanford University, Computer Science Department, 353 Serra Street, Stanford, CA 94305

Abstract

We study a novel model of bounded rationality, one based on an explicit modeling of both abstraction and approximation. Given a zero-sum game and an abstraction for player 2 (that is, a restriction of her strategy space to some compact, convex subset of the simplex), the *exploitability graph (EG)* is a subregion of $\mathbb{R}^2$. Each point of the EG corresponds to one or more strategies of player 1, and captures her exploitability both by an abstracted player 2 and by an unabstracted one. Thus the EG captures the risk incurred by player 1 when reasoning about the assumed opponent rather than the actual one, both at equilibrium and away from it. Our goal in this paper is to characterize the shape of the EG. We first observe that (topologically speaking) the EG is connected, but not necessarily simply connected. We then present our main results, which concern the lowermost points of the EG (the most desirable points from player 1's point of view). We show that the following holds of this "lower envelope": (a) the globally minimum points are contiguous; (b) to the left of the minimum, the points form a convex, monotonically decreasing function, and (c) to the right, they form a strictly monotonically (but not necessarily convex) increasing function. We also discuss what happens to the EG when one varies the abstraction.

Keywords: zero-sum games, abstraction, approximation, bounded rationality, exploitability graph

1. Introduction

The idealizations built into standard game theoretic analysis – in particular the applications of Nash equilibrium and its variants – have been a consistent thorn in the thigh of the field. In real life, agents do not have perfect reasoning capabilities, do not model each other with infinite recursion, and so on. The

*Corresponding author

Email addresses: amarple@stanford.edu (Adrian Marple), shoham@cs.stanford.edu (Yoav Shoham)

[1]This material is based upon work supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) under grant number MURI FA9550-10-1-0573.

literature on so-called *bounded rationality* (cf. [1, 2]), whose aim has been to help bridge the gap between theory and reality, has largely adopted models from computer science. Most commonly, it has used two modeling tools: (a) restricting the strategy space of agents by assuming that they are either computed or implemented by automata within a certain complexity class, or (b) explicitly modeling the cost of computation as a function of the automaton size, and factoring the cost into the agent's utility function.

We pursue a different direction for modeling the fallibility of agents, which does not appeal to complexity or automata-theoretic concepts, nor indeed directly to any computational notions.[2] When the games are too large to reason about, one can abstract one or more agents by projecting their strategy space into a smaller space – any compact, convex subregion of the simplex. And if computing the exact NE in even that smaller game is too expensive, one can approximate the computation. In this paper we provide a theoretical analysis of the interaction between such abstraction and approximation in games. We start with the simplest possible class – zero-sum games in normal form. [3] defined the notion of *exploitability* for this class of games; the exploitability of strategy σ_1 of player 1 is the difference between the value of the game (that is, the payoff to player 1 in equilibrium), and the payoff to player 1 under player 2's best response to σ_1 (clearly, exploitability is always non-negative, and zero only when σ_1 is an equilibrium strategy). Given this definition, any strategy σ_1 of player 1 and abstraction α of player 2's strategy space map to a pair $(x, y) \in \mathbb{R}^2$, where x is σ_1's exploitability by an unconstrained player 2, and y is its exploitability when player 2 is constrained to play in α. We define the *exploitability graph (EG)*, a subregion of $\mathbb{R}^2$, as the image of this mapping.

The EG is one way of capturing the interaction between abstraction and approximation in games. Consider a zero-sum game from player 1's point of view. Assume that she cannot reason about the full strategy space of the game, and instead simplifies it by abstracting the strategy space of the opponent. Which of her strategies are most desirable for her in this abstracted game? To get some intuition for the issues, consider Figure 1, which shows the EG of a simple zero-sum game (the exact game is immaterial here). First note that, from player 1's point of view, the lower envelope (that is, the lowermost points) of the EG dominate the points above them; for a given level of exploitability in the abstracted game, they are the least exploitable strategies in the full game. Also note that the NE strategy in the abstracted game – the leftmost point of the EG, lying on the y-axis – is more exploitable in the full game than some non-equilibrium strategies. This makes perfect sense, since the optimization for the abstract game is not the optimization one would make in the the full game.

In fact, we see in this example that there is a unique global minimum of the lower envelope. This suggests that, under the given abstraction, player 1 should compute the NE only to a certain precision, beyond which the precision

[2]We recognize the irony of having two computer scientists veer away from computational models.

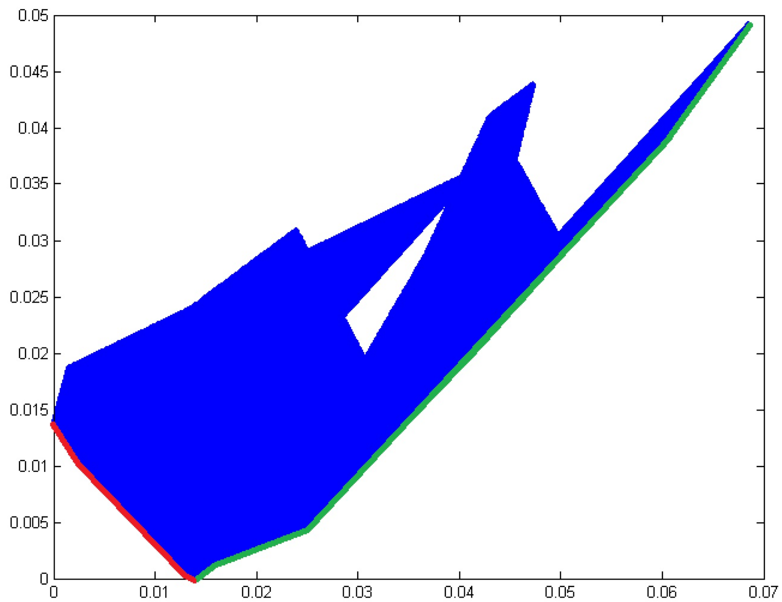

Figure 1: An example exploitability graph (EG). The x-axis displays abstract exploitability and the y-axis displays full game exploitability. Note that exploitability is relative to the game value, and the abstract game and full game have different values; hence the EG does not lie entirely above the diagonal. The red line represents the left lower envelope and the green represents the right lower envelope. (Although it is immaterial here, for completeness we mention that the full game generating this EG was a 3x12 normal-form zero-sum game with values drawn randomly from $[0, 1)$, and the abstract game eliminated an arbitrarily chosen half of player 2's pure strategies.)

will be counter-productive.[3] Indeed, this "over-fitting" phenomenon has been observed in practice. After years of poker programs performing poorly against even mediocre human players, today the best programs playing Limit Texas Hold'em hold their own even against expert players (cf. [4]). The change took place when the programs began to abstract the games by collapsing information sets (that is, ignoring distinctions among certain cards), and computing the approximate BNE of the simplified game using heuristic algorithms where the over-fitting phenomenon was observed experimentally [4]. Our goal in this paper is to understand the mathematical structure of the EG in general, and the connection to the poker literature is for now only suggestive. Still, the poker example (among others) suggests that this approach to modeling bounded ra-

[3]Of course, any actual procedure to progressively approximate the NE will not necessarily traverse the lower enveloper, although in practice – for example, in the poker example discussed in this paragrpah – such procedures tend to follow its general contour.

tionality holds the promise of connecting theory and practice.

Our formal results essentially show that the properties of the EG seen in Figure 1 hold in general. To begin with, we observe that the EG is always connected, but not necessarily simply connected. Our main results concern the lower envelope of the EG. We show that (a) the globally minimum points are contiguous, (b) to the left of the minimum the points form a convex, monotonically decreasing function, and (c) to the right they form a strictly monotonically (but not necessarily convex) increasing function.

At this point questions such as the following may have come to mind: Why do we restrict ourselves to comparing a reference abstraction of player 2 to an unabstracted player 2, as opposed to some other abstraction (coarser, finer, or non-comparable) of that player? Or, what happens to the EG as we vary the reference abstraction of player 2? Or, should we also consider abstractions of player 1's strategy space? After all, in general player 1's full strategy space is as hard to reason about as player 2's. Or, what about games not in normal formal, or non-zero-sum games?

We will address all these questions, but it will be easier to do so after we've developed the technical material, and so we return to them in sections 5 and 6.

The rest of the paper is organized as follows: In section 2 we cover relevant background material; section 3 defines the exploitability graph and related constructs, and discuss some basic topological properties of exploitability graphs; in section 4 we prove our main results, concerning the exploitability graph; in section 5 we discuss some findings regarding the effects of varying player 2's level of abstraction on the EG; and section 6 summarizes the results and discusses possible extensions.

2. Background

We will be primarily concerned with zero-sum normal form games, although in section 7 we discuss the possible extensions to other classes of games. We assume familiarity with the relevant concepts and notation. The following sections introduce concepts that are not (or not completely) standard.

2.1. Abstraction

Definition 2.1. A *strategy abstraction* is a restriction of a player's strategy space to a compact, convex subspace of her original strategy space.

The *null abstraction* leaves the strategy space unchanged. An *abstract game* is a game derived from applying a strategy abstraction to all players. Note that abstract games clearly retain standard game properties such as existence of NE by standard fixed point arguments.

Note: Typically, abstractions reduce the strategy space in specific ways (for example, by removing certain pure strategies and the mixed strategies that include them), but it is important for our development to consider the broader class of abstractions.

2.2. Exploitability

The final piece of background information needed is the notion of exploitability. Exploitability, introduced by Waugh et al., is a simple construct that evaluates the fitness of a particular strategy in the context of a zero-sum two-person game [3]. Formally, we define player 1's exploitability as follows:

Definition 2.2 (Exploitability). Let b_2 denote player 2's value of her best response and let v^* be the player 1 game value. The *exploitability* of a player 1 strategy, σ_1, is defined to be

$$\epsilon_1(\sigma_1) = b_2(\sigma_1) + v^*$$

Exploitability can be thought of as the degree to which player 2 can force a given player 1 strategy away from the equilibrium game value. Note that exploitability is always non-negative and is equal to zero if and only if the strategy in question is an equilibrium strategy.

3. Exploitability Graph

Until we generalize the discussion in sections 6 and 7, we will be comparing a pair of abstractions of a (two-person) zero-sum game: One in which neither player is abstracted (called the *full game*) and one in which player 2 (often referred to as the opponent) has some abstraction (the *abstract game*). We will use ϵ_a and ϵ_f to to denote player 1's exploitability function in the abstract and full games, respectively. Player 1's strategy space (identical in both games) will be denoted Σ_1.

The exploitability graph (EG) is a useful tool for viewing the effect of both abstraction (in this case a single pair of abstractions) and approximation (exploitability). Simply put, the EG plots two exploitability values, abstract and full, associated with every strategy. More formally

Definition 3.1 (Exploitability Graph). The *exploitability graph* is the image of the following function, $E : \Sigma_1 \longrightarrow \mathbb{R}^2$:

$$E : \sigma_1 \longmapsto (\epsilon_a(\sigma_1), \epsilon_f(\sigma_1))$$

The exploitability graph is a surprisingly rich and complex object and part of its strength as a tool is how well it is represented visually. For example, the EG in Figure 1 appears to be connected, but not simply connected. Is this accidental?

It turns out these are indeed general properties of EGs. To see that connectedness always holds, note that exploitability is continuous if and only if the player 2 best response value function, b_2, is continuous (since they differ only by a constant). However, the latter continuity is obvious because it is the maximum of continuous functions (over a compact domain). Now by exploitability's contintuty, the function E used to define the EG is continuous. Hence, because

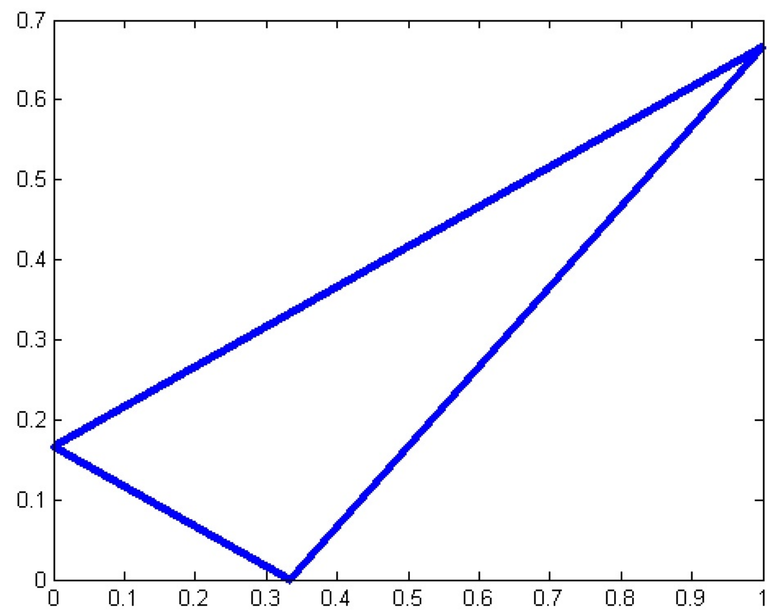

Figure 2: The exploitability graph of a game that is not simply connected.

the player 1 strategy space is connected (due to convexity) and the image of a connected set under a continuous map is connected, the EG is connected.

Conversely, to see that there exist exploitability graphs that are not simply connected, an example suffices. A simple such example is seen in Figure 2. This EG corresponds to the game described by the following player 1 payoff matrix (note that this defines player 2's payoffs as well since it is a zero-sum game):

-1	1	0
1	-1	-1

In this game, the column player (player 2) has all three actions available when unabstracted, but when abstracted player 2 is restricted to the left two actions. To see that this forms a loop, consider that when player 1 plays a pure strategy he receives a payoff of -1 (after a player 2 best response) therefore the endpoints of the strategy space clearly intersect. However, the two games have different equilibrium strategies. It is simple to verify that this causes the exploitability to take the form of a non-trivial loop as can be seen in Figure 2.

While such general properties of the EG are interesting, as was discussed in the introduction, most interesting is its lower envelope. It is defined as follows:

Definition 3.2 (Lower Envelope). Let the domain of the lower envelope be

$D = [0, \max(\epsilon_a(\Sigma_1))]$. The *lower envelope* of the EG the following function, $\ell : D \mapsto \mathbb{R}$:

$$\ell(x) = \min_{\sigma_1 \in \Sigma_1 | \epsilon_a(\sigma_1) = x} \epsilon_f(\sigma_1)$$

We should note that ℓ is in fact well defined even though *a priori* we don't know whether $\{\sigma_1 \in \Sigma_1 | \epsilon_a(\sigma_1) = x\}$ is non-empty for all values between 0 and $\max(\epsilon_a(\Sigma_1))$. However, as will be shown in section **??** exploitability is continuous and thus by the intermediate value theorem ℓ is well defined.

In addition to the lower envelope, we define the left and right envelopes.

Definition 3.3 (Left and Right Lower Envelopes). Let z be the largest value that minimizes ℓ, i.e.

$$z = \max(\arg\min_x \ell(x))$$

Then the *left and right lower envelopes* respectively are defined by:

$$\ell^- = \ell|_{[0,z]}$$
$$\ell^+ = \ell|_{[z,\max(\epsilon_a(\Sigma_1))]}$$

Thus the left lower envelope is the lower envelope to the left of z and the right lower envelope is the lower envelope to the right of z.

Again a visualization of the lower envelope is useful. Figure 1 shows the left and right lower envelopes in red and green, respectively.

4. Convexity and Monotonicity Theorems

Our two central theorems concern the shape of the lower envelope:

Theorem 4.1 (Convexity of the Left Lower Envelope). *ℓ^- is a convex, monotonically decreasing function.*

Theorem 4.2 (Monotonicity of the Right Lower Envelope). *ℓ^+ is a strictly monotonically increasing function.*

In the remainder of this section we prove these two theorems.

4.1. Exploitability-Bounded Strategy Restriction

Necessary for the proofs of these two theorems is the notion of a exploitability-bounded strategy restriction (EBSR).

Definition 4.1 (EBSR). The *exploitability-bounded strategy restriction* for player i as a function of $x \geq 0$ is defined to be

$$\Sigma_i(x) = \{\sigma_i \in \Sigma_i | \epsilon_i(\sigma_i) \leq x\}$$

Thus a EBSR is simply the subset of the player 1's strategy space that is exploitable up to a particular value. Similar to our definitions of ϵ_a and ϵ_f, let $\Sigma_a(x)$ and $\Sigma_f(x)$ denote the player 1 EBSRs of the abstract and full games respectively.

Lemma 4.1. *Given that Σ_1 is convex, Σ_2 is compact, and $u(\sigma_1, \sigma_2)$ is linear with respect to σ_1 (i.e. $u(\alpha\sigma_1 + (1-\alpha)\sigma_1', \sigma_2) = \alpha u(\sigma_1, \sigma_2) + (1-\alpha)u(\sigma_1', \sigma_2)$) and continuous with respect to σ_2. Then player 2's best response value function, $b_2(\sigma_1) = \max_{\sigma_2}(-u(\sigma_1, \sigma_2))$, is convex.*[4]

Proof. By the compactness of Σ_2 and the continuity of u, b_2 is well defined. Let $0 \le \alpha \le 1$. The following series of equalities and inequalities proves that b_2 is convex

$$\begin{aligned}
b_2(\alpha\sigma_1 + (1-\alpha)\sigma_1') &= \max_{\sigma_2}(-u(\alpha\sigma_1 + (1-\alpha)\sigma_1', \sigma_2)) \\
&= \max_{\sigma_2}(\alpha(-u(\sigma_1, \sigma_2)) + (1-\alpha)(-u(\sigma_1', \sigma_2))) \\
&\le \alpha \max_{\sigma_2}(-u(\sigma_1, \sigma_2)) + (1-\alpha) \max_{\sigma_2}(-u(\sigma_1', \sigma_2)) \\
&= \alpha b_2(\sigma_1) + (1-\alpha) b_2(\sigma_1')
\end{aligned}$$

Lemma 4.2. *If $b_2(\sigma_1)$ is a convex function then $\Sigma_1(x)$ is convex for all $x \ge 0$.*

Proof. Let $x \ge 0$. $\Sigma_1(x)$ is guaranteed to be non-empty since there exists a strategy that achieves an exploitability of 0. Because of this we can choose two strategies, $\sigma_1, \sigma_1' \in \Sigma_1(x)$. Let $0 \le \alpha \le 1$. Now consider the exploitability of a convex combination of these two strategies.

$$\begin{aligned}
\epsilon_1(\alpha\sigma_1 + (1-\alpha)\sigma_1') &= b_2(\alpha\sigma_1 + (1-\alpha)\sigma_1') + v^* \\
&\le \alpha b_2(\sigma_1) + (1-\alpha) b_2(\sigma_1') + v^* \\
&= \alpha(b_2(\sigma_1) + v^*) + (1-\alpha)(b_2(\sigma_1') + v^*) \\
&= \alpha\epsilon_1(\sigma_1) + (1-\alpha)\epsilon_1(\sigma_1') \\
&\le \alpha x + (1-\alpha)x \\
&= x
\end{aligned}$$

The two inequality steps hold by the assumption that b_2 is convex and the definition of EBSR respectively. This shows that a convex combination of strategies in $\Sigma_1(x)$ is in $\Sigma_1(x)$, proving that $\Sigma_1(x)$ is convex for all $x \ge 0$.

4.2. Proof of the Convexity Theorem

Proof (of Theorem 4.1). Assume ℓ^- is not convex. In this case, there exists a line ($y = mx + b$) that is less than or equal to ℓ^- but intersects ℓ^- at two

[4]It should also be noted, although we will provide no formal proof, that the various maximums and minimums used in the following material exist by the compactness of the strategy spaces and continuity of the functions involved.

points, x_1, x_2, such that $x_1 < x_2$ and ℓ^- is strictly above the line at the average of the two points, i.e.

$$m\frac{x_1 + x_2}{2} + b < \ell^-\left(\frac{(x_1 + x_2)}{2}\right).$$

First, we can notice that m must be non-positive. Otherwise if x_1 strictly minimizes ℓ^-, then $\ell^-(x_1) < \min_x(\ell^-(x))$ (a contradiction). But if x_1 does not strictly minimize, consider the rightmost point, z. First remember that z minimizes ℓ^- and by how x_1 is picked $x_1 < x_2 \leq z$. These two points combined with the assumption that m is positive yields $\ell^-(z) \leq \ell^-(x_1) = mx_1+b < mz+b$ contradicting that ℓ^- is always less that or equal to the line $y = mx + b$.

Now consider the game where player 2's strategy space is replaced by

$$\Sigma_2' = \Sigma_2^a \times \Sigma_2$$

(i.e. the cross product of player 2's full game and abstract strategy spaces) and player 2's payoff function is replaced by

$$u_2' : \Sigma_2' \longrightarrow \mathbb{R}$$
$$u_2'(\sigma_2, \sigma_2') = c \cdot u_2(\sigma_2) + u_2(\sigma_2')$$

It should be noted that this still satisfies the general class of games that we are using. Σ_2' is clearly convex and compact since both Σ_2^a and Σ_2 are and u_2' is just as clearly linear with respect to player 2's new strategy space since it is the linear combination of two similarly linear functions. Now let $\Sigma_c(x)$ denote $\Sigma_1(x)$ for this game. An important point to note is that so long as c is greater than or equal to 0, player 2's best response value to a player 1 strategy, σ_1 is given by $c \cdot b_2^a(\sigma_1) + b_2^f(\sigma_1)$ and thus player 1's exploitability in this game, denoted $\epsilon_c(\sigma_1)$, is given by $c\epsilon_a(\sigma_1) + \epsilon_f(\sigma_1) - k$.

Considering again the aforementioned line, consider Σ_{-m} (Σ_c with c replaced by $-m$). Also choose $\sigma_1, \sigma_1' \in \Sigma_1$ such that $\epsilon_a(\sigma_1) = x_1, \epsilon_f(\sigma_1) = \ell(x_1), \epsilon_a(\sigma_1') = x_2$, and $\epsilon_f(\sigma_1') = \ell(x_2)$. Notice that because σ_1 and σ_1' both map to the line $y = mx + b$, $-m\epsilon_a(\sigma_1) + \epsilon_f(\sigma_1) = b$ and $-m\epsilon_a(\sigma_1') + \epsilon_f(\sigma_1') = b$. Therefore we have $\epsilon_{-m}(\sigma) = \epsilon_{-m}(\sigma') = b - k$ (again ϵ_{-m} is ϵ_c with c replaced by $-m$). Furthermore, for a strategy to attain a lower value of ϵ_{-m} it must fall below the line $y = mx + b$. However, because ℓ^- lies entirely at or above this line, the entire exploitability graph lies at or above this line and hence there are no strategies that have an ϵ_{-m} value less than $b - k$. From this we can conclude that $k = b$ and $\sigma_1, \sigma_1' \in \Sigma_{-m}(0)$.

If we look at the function $\alpha \mapsto \epsilon_a(\alpha\sigma_a + (1-\alpha)\sigma_1')$, by the intermediate value theorem, there must be an α that produces a strategy that attains an abstract exploitability value of $\frac{x_1+x_2}{2}$. Moreover, because this function uses only convex combinations of σ_1 and σ_1' the strategy that attains this average value is also in $\Sigma_{-m}(0)$ by the convexity of $\Sigma_{-m}(0)$. However, this means that there is a strategy with abstract exploitability $\frac{x_1+x_2}{2}$ that lies on the line $y = mx + b$, implying that ℓ^- is equal to the line at $\frac{x_1+x_2}{2}$, which contradicts how $y = mx+b$

was chosen. Thus the assumption that ℓ^- is not convex is false. Because ℓ^- is convex and it attains its minimum value at its rightmost point, it is also a monotonically decreasing function.

4.3. Proof of the Monotonicity Theorem

In order to prove the monotonicity of the right lower envelope (ℓ^+), an "almost inverse" function to ℓ^+, denoted h, needs to be defined. h represents the rightmost point (most abstract exploitable point) on the exploitability graph below a certain full exploitability value, and is formally defined as

Definition 4.2 (Dual of the Left Lower Envelope).

$$h(x) = \max(\epsilon_a(\Sigma_f(x)))$$

As can be seen in Figure 3, h can be thought of as an inverse function of ℓ^+, with certain discontinuities of ℓ^+ filled in. That is to say $\ell^+(h(x)) = x$ so long as ℓ^+ actually achieves the value of x somewhere. It is important to see that if ℓ^+ ever fails to increase, this causes h to jump in a discontinuous manner as shown in Figure 3. The proof of Theorem 4.2 hinges on this fact and is formally defined and proven in the following lemma:

Lemma 4.3. *If ℓ^+ is not strictly monotonic, h has a left discontinuity*[5]*.*

Proof. If ℓ^+ is not strictly monotonic, then it will have a local minimum (a point $c \neq z$ such that there exists an $\epsilon > 0$ such that $|x - c| < \epsilon$ implies $\ell^+(x) > \ell^+(c)$). Now consider the greatest such c and a strategy, σ_1, that maps to $(c, \ell^+(c))$. Note that $h(\ell^+(c)) = c$. If this were otherwise, then there would be a strategy that mapped to (α, β) where $\alpha > c$ and $\beta \leq \ell^+(c)$. For this to happen ℓ^+ would have to decrease after c implying there would be a local minimum in ℓ^+ somewhere after c. Similarly, by the fact that c is the last local minimum and by the definition of ℓ^+ there exist no strategies that map to (α, β) where $\alpha > c - \epsilon$ and $\beta < \ell^+(c)$. In other words, for all $\delta > 0$, $|h(\ell^+(c) - \delta) - c| \geq \epsilon$. This is precisely the contrapositive statement of left continuity for h, thus if ℓ^+ is not strictly monotonic h is left discontinuous at c.

There are a few more important facts that are pivotal in the proof of Theorem 4.2. First, because both full and abstract exploitabilities are continuous (see section **??** for proof) and operate over a compact space (Σ_1) they are not only continuous but uniformly continuous. Second, h is a monotonically increasing (not necessarily strictly increasing) function. This is trivial to show since the strategy that maximized exploitability in $\Sigma_f(x)$ is still an element of $\Sigma_f(x+\delta)$ for $\delta > 0$. Finally, the EBSR, $(\Sigma_f(x))$ itself has an important continuity property that is captured in the following lemma:

[5] Left continuity is defined by: $\forall \epsilon > 0, c > 0$ there exists $\delta > 0$ such that if $c - \delta < x < c$ then $|f(x) - f(c)| < \epsilon$

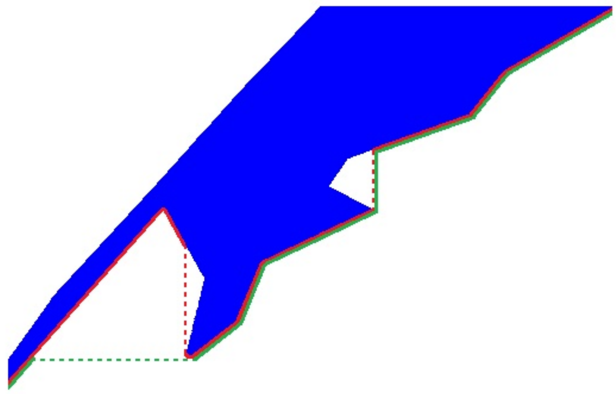

Figure 3: A depiction of the left lower envelope, ℓ^+, in red and its dual function, h, in green of a fictitious EG. Discontinuities are depicted by dotted lines. This depiction shows a case when the left lower envelope decreases causing a discontinuity in its dual (although by theorem 4.2 this cannot happen in a real EG). Also, notice that ℓ^+ is a function with respect to the x-axis and h is a function with respect to the y-axis.

Lemma 4.4. *Let $x > 0$. For all $\epsilon > 0$, there exists a $\delta > 0$ such that for all points $\sigma_1 \in \Sigma_f(x)$ there exists a point $\sigma_1' \in \Sigma_f(x - \delta)$ such that the distance between σ_1 and σ_1' is less than ϵ.*

Proof. Let $\epsilon > 0$ and let $\delta_0 > 0, \delta_0 < x$. Consider a point $\sigma_1 \in \Sigma_f(x)$. Now let σ_1' be one of the closest points to σ_1 that is in $\Sigma_f(x-\delta_0)$. Let $\alpha \in (0, 1]$ and let $\sigma_1^* = \sigma_1'\alpha + \sigma_1(1-\alpha)$. Because $\Sigma_f(x)$ is convex σ_1^* is still in $\Sigma_f(x)$. Furthermore, because σ_1' satisfies $A\sigma_1' \geq v^A - (x-\delta_0)$ and σ_1 satisfies $A\sigma_1' \geq v^A - x$, σ_1^* satisfies $A\sigma_1^* \geq v^A - (x - \alpha\delta_0)$. Thus σ_1^* is an element of $\Sigma_f(x - \alpha\delta_0)$. Also, if d_0 is the distance between σ_1 and σ_1', it follows from the definition of σ_1^* that the distance between σ_1 and σ_1^* is αd_0.

Let d be the maximum distance between a point in $\Sigma_f(x)$ and its closest point in $\Sigma_f(x - \delta_0)$ (this exist by compactness of $\Sigma_1(x)$). Now choose α such that $\alpha d < \epsilon$ and choose δ to be $\alpha\delta_0$. For every point $\sigma_1 \in \Sigma_f(x)$ by the above arguments there will be a point in $\Sigma_f(x - \delta)$ that is at most αd away from σ_1. Thus we have shown the existence of a δ satisfying the lemma.

We now have all the tools necessary to complete the proof that ℓ^+ is monotonic.

Proof (of Theorem 4.2). Let $\epsilon > 0$ and $x > 0$. By the uniform continuity of exploitability we can choose a $\delta_1 > 0$ such that for all $\sigma_1 \in \Sigma_1, |\sigma_1 - \sigma_1'| < \delta_1$ implies $|\epsilon_a(\sigma_1) - \epsilon_a(\sigma_1')| < \epsilon$. Now by Lemma 4.4 we can choose a $\delta > 0$ such that for all $\sigma_1 \in \Sigma_f(x)$ there exists a point $\sigma_1' \in \Sigma_f(x-\delta)$ such that $|\sigma_1 - \sigma_1'| < \delta_1$. Together these imply that the exploitability of a point in $\Sigma_f(x)$ must differ less than ϵ from the exploitability of some point in $\Sigma_f(x - \delta)$. Specifically, if we take the max of all the exploitabilities of points in $\Sigma_f(x)$, (i.e. $h(x)$), it must differ less than ϵ for the exploitability of some point in $\Sigma_f(x-\delta)$. By the monotonicity of h this point must have exploitability less than $h(x)$ and thus the max of all the exploitabilities of points in $\Sigma_f(x-\delta)$, (i.e. $h(x-\delta)$), must also differ from $h(x)$ by less than ϵ. This is equivalent to saying that we have found a δ such that $|h(x-\delta) - h(x)| < \epsilon$. Furthermore, because h is monotonically increasing, $x' \in (x-\delta, x)$ implies $|h(x') - h(x)| \leq |h(x-\delta) - h(x)|$. Altogether this is sufficient to show that h is left continuous. By the contrapositive of Lemma 4.3, this means that ℓ^+ is strictly monotonically increasing.

5. Varying Player 2's Abstraction

So far we have only considered comparing one reference abstraction of player 2 to the null abstraction, but we have left opened what might happen when we consider a sequence of such reference abstractions. If the abstractions are unrelated to each other, one cannot hope to discover any pattern as one varies the abstraction; each abstraction gives rise to an arbitrarily different game. So we will consider chains in which the abstraction level increases monotonically:

Definition 5.1. Hold some normal form game fixed. A *chain of abstractions (of player 2)* of this game is a finite list of strategy abstractions $\alpha_0, \ldots \alpha_n$ for player 2 such that for all $i \in \{1, \ldots, n\}$ α_{i-1} is a coarser abstraction than α_i (i.e. the player 2 strategy space in α_{i-1} is a subset of the player 2 strategy space in α_i).

In fact, this restriction alone is not sufficient to give rise to any regularity. To see this imagine that we want the next abstraction to effectively be a particular game with the same player 1 strategy space. We can simply subtract a constant factor from this game until all of its values strictly dominate the value of the previous game (from player 2's point of view). If these adjusted strategies are now introduced as the refinement of an abstraction, the previous game will cease to be relevant. Therefore, if we allow for this sort of abstraction then we have little hope of seeing patterns in the EG's of a chain of abstractions. And so we consider one additional restriction, and introduce a "grain of truth" requirement.[6]

Definition 5.2 (Grain of Truth). Given a game and an abstraction α of this game. α satisfies the *grain of truth* requirement if it contains a strategy that is in player 2's support of an equilibrium of the full game. Likewise, let $\alpha_0, \ldots, \alpha_n$ be a chain of abstractions for this game. This chain satisfies the *grain of truth* requirement if α_0 satisfies the grain of truth requirement.

Note that by the definition of a chain of abstractions, if the bottommost abstraction α_0 has a grain of truth then so do all abstraction of the chain.

[6] To avoid confusion, we should mention that this is different from the "grain of truth" condition introduced by Kalai and Lehrer in [5].

Armed with this grain of truth requirement we can finally find some pattern in chains of abstractions.

Proposition 5.1. *Given a game and a chain of abstractions $\alpha_0, \ldots, \alpha_i, \ldots, \alpha_n$ that satisfies the grain of truth requirement. Then the EG corresponding to each α_i (versus the full game) has a single minimum, and that minimum moves monotonically to the left as one increases abstraction index i.*

Proof. Recall the definition of exploitability.

$$\epsilon_i(\sigma_1) = b_2(\sigma_1) + v_i^*$$

where ϵ_i denotes the exploitability of abstraction α_i and v_i^* represents the player 1 game value of abstraction α_i. Recall that a minimum of the EG for any abstraction must corresponed to an equilibrium of the full game. However, because of the grain of truth requirement each abstraction must have a strategy in the support of a full game equilibrium. Because we are dealing with zero-sum games this means that the value of any player 2 best response to any full game equilibrium is the same. Namely, we have for any player 1 strategy of a full game Nash equilibrium σ_1^*

$$\epsilon_i(\sigma_1^*) = v_i^* - v^*$$

Therefore all minima (which are necessarily Nash equilibria of the full game) have the same abstract exploitability which implies that the EG of each abstraction has a single minimum. Furthermore, it is simple to see that v_i^* can only decrease as player 2 gains more strategies (i.e. as i increases). Thus this minimum point will move monotonically to the left with increasing abstraction index.

6. Discussion

We have defined the exploitability graph (EG) as a new formal object with which to study bounded rationality, and provided some basic characterizations of it. Our main results concerned its lower envelope, which we showed to have a specific structure (a global minimum point or region; to its left a monotonically decreasing and convex region, and to its right a monotonically strictly increasing region).

In the introduction we acknowledged certain questions that might arise about the model, and promised to address them after the formal development. One of them – the effect on the EG of varying the reference abstraction of player 2 – was addressed in section 5. We address the others below.

- *You analyse the risk of under-representing an opponent's strength; what about over-representing it, or incorrectly representing it in some other way?* It is true that we defined the EG as comparing the exploitability by a full opponent versus a more limited (or abstracted) one, and indeed our motivation stemmed from situations (such as computational poker) in which opponent's actual capabilities are modelled in curtailed form. But our results actual hold in much more general settings. In particular, they hold for any arbitrary pair of abstractions of player 2. This means that the abstraction that we have been taking to be the null abstraction (a full powered opponent) can in general be coarser than, finer than, or non-comparable to the reference abstraction. For example, if we switch the roles the two abstractions play we will have effectively flipped the axes of the EG. Because the results hold for an arbitrary pair, they will still hold in this case. Thus, what was previously the "left" envelope will now be the lower envelope and all the properties we prove will hold.

∗ *What about abstracting player 1's own strategy space, which in general is as hard to model as player 2's?* Indeed we assume that when reasoning about the game, player 1 would abstract her own strategy as well. But since that abstraction is held fixed as one considers both the abstract opponent and the unconstrained one, it plays no role in our analysis.

∗ *But still, what happens to the EG as you vary player 1's abstraction, akin to your analysis for player 2 in section 5*? That is a worthwhile question which we don't study in this paper.

∗ *What about an interactive analysis, in which both players simultaneously consider abstractions of the opponent, and perhaps their own?* We do not consider this situation at all. In some sense it undercuts the motivation for our approach. At least a straightforward approach, which would model all possible abstractions by both players and seek an equilibrium between the abstractions, would blow up the strategy space rather than contract it. But it is an intellectually interesting direction to pursue, and potentially also useful from the pragmatic standpoint.

∗ *What about games not in normal form?* All games that have equivalent normal-form games are almost immediately covered by our analysis. In particular, extensive-form games with mixed strategies are covered, as are games with behavioral strategies with perfect recall (since those have equivalent games with mixed strategies [6]). One can extend the treatment to games of imperfect recall with considerable additional conceptual and notational development, which we omit.

∗ *What about non-zero-sum games?* One can generalize the definition of exploitability to general-sum games, and extend our results. But this extension involves no novel conceptual or technical ideas, so we omit it too.

To conclude, as discussed in the introduction, reducing the gap between the mathematical models of game theory and real situations is essential for the long-term success of the field, as it is for any natural or social science. We hope that the mathematical study of the exploitability graph, started here, will help narrow this gap.

References

[1] A. Rubinstein, Modeling Bounded Rationality, 1998.

[2] A. Neyman, Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma, Economics Letters 19 (1985) 227–229.

[3] K. Waugh, D. Schnizlein, M. Bowling, D. Szafron, Abstraction pathologies in extensive games, in: Proceedings of the 8th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), 2009.

[4] M. Johanson, K. Waugh, M. Bowling, M. Zinkevich, Accelerating best response calculation in large extensive games, in: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI-11), 2011.

[5] E. Kalai, E. Lehrer, Rational learning leads to Nash equilibrium, Econometrica 61.

[6] H. Kuhn, Entensive Games and the Problem of Information, Vol. 2 of Contributions to the Theory of Games, Princeton University Press, 1953, pp. pp. 193–216.