Iterated Regret Minimization: A New Solution Concept

Joseph Y. Halpern and **Rafael Pass**
Cornell University
{halpern,rafael}@cs.cornell.edu

Abstract

For some well-known games, such as the Traveler's Dilemma or the Centipede Game, traditional game-theoretic solution concepts—most notably Nash equilibrium—predict outcomes that are not consistent with empirical observations. We introduce a new solution concept, *iterated regret minimization*, which exhibits the same qualitative behavior as that observed in experiments in many games of interest, including Traveler's Dilemma, the Centipede Game, Nash bargaining, and Bertrand competition. As the name suggests, iterated regret minimization involves the iterated deletion of strategies that do not minimize regret.

1 Introduction

Perhaps the most common solution concept considered in game theory is Nash equilibrium. While many other solution concepts have been considered (see [Osborne and Rubinstein, Osborne and Rubinstein1994] for an overview), there are a number of well-known games where none of them seems appropriate. Consider the well-known *Traveler's Dilemma* [Basu, Basu1994; Basu, Basu2007]. Suppose that two travelers have identical luggage, for which they both paid the same price. Their luggage is damaged (in an identical way) by an airline. The airline offers to recompense them for their luggage. They may ask for any dollar amount between \$2 and \$100. There is only one catch. If they ask for the same amount, then that is what they will both receive. However, if they ask for different amounts—say one asks for \$m and the other for \$m′, with $m < m′$—then whoever asks for \$m (the lower amount) will get \$$(m+p)$, while the other traveler will get \$$(m-p)$, where p can be viewed as a reward for the person who asked for the lower amount, and a penalty for the person who asked for the higher amount.

It seems at first blush that both travelers should ask for \$100, the maximum amount, for then they will both get that. However, as long as $p > 1$, one of them might then realize that he is actually better off asking for \$99 if the other traveler asks for \$100, since he then gets \$101. In fact, \$99 *weakly dominates* \$100, in that no matter what Traveler 1 asks for, Traveler 2 is always at least as well off asking for \$99 than \$100, and in one case (if Traveler 2 asks for \$100) Traveler

1 is strictly better off asking for \$99. Thus, it seems we can eliminate 100 as an amount to ask for. However, if we eliminate 100, a similar argument shows that 98 weakly dominates 99! And once we eliminate 99, then 97 weakly dominates 98. Continuing this argument, both travelers end up asking for \$2! In fact, it is easy to see that (2,2) is the only Nash equilibrium. With any other pair of requests, at least one of the travelers would want to change his request if he knew what the other traveler was asking. Since (2,2) is the only Nash equilibrium, it is also the only sequential and perfect equilibrium. Moreover, it is the only rationalizable strategy profile; indeed (once we also consider mixed strategies), (2,2) is the only strategy that survives iterated deletion of strongly dominated strategies. (It is not necessary to understand these solution concepts in detail; the only point we are trying make here is that all standard solution concepts lead to (2,2).)

This seems like a strange result. It seems that no reasonable person—even a game theorist!—would ever play 2. Indeed, when the Traveler's Dilemma was empirically tested among game theorists (with $p = 2$) they typically did not play anywhere close to 2. Becker, Carter, and Naeve [2005] asked members of the Game Theory Society (presumably, all experts in game theory) to submit a strategy for the game. Fifty-one of them did so. Of the 45 that submitted pure strategies, 33 submitted a strategy of 95 or higher, and 38 submitted a strategy of 90 or higher; only 3 submitted the "recommended" strategy of 2. The strategy that performed best (in pairwise matchups against all submitted strategies) was 97, which had an average payoff of \$85.09. The worst average payoff went to those who played 2; it was only \$3.92.

Another sequence of experiments by Capra et al. [1999] showed, among other things, that this result was quite sensitive to the choice of p. For low values of p, people tended to play high values, and keep playing them when the game was repeated. By way of contrast, for high values of p, people started much lower, and converged to playing 2 after a few rounds of repeated play. The standard solution concepts (Nash equilibrium, rationalizability, etc.) are all insensitive to the choice of p; for example, (2,2) is the only Nash equilibrium for all choices of $p > 1$.

We introduce a new solution concept, *iterated regret minimization*, which has the same qualitative behavior as that observed in the experiments, not just in Traveler's Dilemma, but in many other games that have proved problematic for Nash

equilibrium, including the Centipede Game, Nash bargaining, and Bertrand competition. We focus on iterated regret minimization in strategic games, and comment on how it can be applied to Bayesian games.

2 Iterated Regret Minimization

We start by providing an informal discussion of iterated regret minimization in strategic (normal-form) games, and applying it to the Traveler's Dilemma; we then give a more formal treatment.

Nash equilibrium implicitly assumes that the players know what strategy the other players are using. (See [Aumann and Brandenburger, Aumann and Brandenburger1995] for a discussion of the knowledge required for Nash equilibrium.) Such knowledge seems unreasonable, especially in one-shot games. Regret minimization is one way of trying to capture the intuition that a player wants to do well no matter what the other players do.

The idea of minimizing regret was introduced (independently) in decision theory by Savage [1951] and Niehans [1948]. To explain how we use it in a game-theoretic context, we first review how it works in a single-agent decision problem. Suppose that an agent chooses an act from a set A of acts. The agent is uncertain as to the true state of the world; there is a set S of possible states. Associated with each state $s \in S$ and act $a \in A$ is the utility $u(a, s)$ of performing act a if s is the true state of the world. For simplicity, we take S and A to be finite here. The idea behind the *minimax regret* rule is to hedge the agent's bets, by doing reasonably well no matter what the actual state is. For each state s, let $u^*(s)$ be the best outcome in state s; that is, $u^*(s) = \max_{a \in A} u(a, s)$. The *regret* of a in state s, denoted $regret_u(a, s)$, is $u^*(s) - u(a, s)$; that is, the regret of a in s is the difference between the utility of the best possible outcome in s and the utility of performing act a in s. Let $regret_u(a) = \max_{s \in S} regret_u(a, s)$. For example, if $regret_u(a) = 2$, then in each state s, the utility of performing a in s is guaranteed to be within 2 of the utility of any act the agent could choose, even if she knew that the actual state was s. The minimax-regret decision rule orders acts by their regret; the "best" act is the one that minimizes regret. Intuitively, this rule is trying to minimize the regret that an agent would feel if she discovered what the situation actually was: the "I wish I had chosen a' instead of a" feeling.

Despite having been used in decision making for over 50 years, up until recently, there was little work on applying regret minimization in the context of game theory. We discuss related work in Section 4; here, we describe our own approach. We start by explaining it in the context of the Traveler's Dilemma and restrict our attention to pure strategies. We take the acts for one player to be that player's pure strategy choices and take the states to be the other player's pure strategy choices. Each act-state pair is then just a strategy profile; the utility of the act-state pair for player i is just the payoff to player i of the strategy profile. Intuitively, each agent is uncertain about what the other agent will do, and tries to choose an act that minimizes his regret, given that uncertainty.

It is easy to see that, if the penalty/reward $2 \leq p \leq 49$,

then the acts that minimize regret are the ones in the interval $[100 - 2p, 100]$; the regret for all these acts is $2p - 1$. For if player 1 asks for an amount $m \in [100 - 2p, 100]$ and player 2 asks for an amount $m' \leq m$, then the payoff to player 1 is at least $m' - p$, compared to the payoff of $m' + p - 1$ (or just m' if $m' = 2$) that is achieved with the best response; thus, the regret is at most $2p - 1$ in this case. If, instead, player 2 asks for $m' > m$, then player 1 gets a payoff of $m + p$, and the best possible payoff in the game is $99 + p$, so his regret is at most $99 - m \leq 2p - 1$. On the other hand, if player 1 chooses $m < 100 - 2p$, then his regret will be $99 - m > 2p - 1$ if player 2 plays 100. On the other hand, if $p \geq 50$, then the unique act that minimizes regret is asking for \$2.

Suppose that $2 \leq p \leq 49$. Applying regret minimization once suggests using a strategy in the interval $[100 - 2p, 100]$. But we can iterate this process. If we assume that both players use a strategy in this interval, then the strategy that minimizes regret is that of asking for \$$(100 - 2p + 1)$. A straightforward check shows that this has regret $2p - 2$; all other strategies have regret $2p - 1$. If $p = 2$, this approach singles out the strategy of asking for \$97, which was found to be the best strategy by Becker, Carter, and Naeve [2005]. As p increases, the act that survives this iterated deletion process goes down, reaching 2 if $p \geq 50$. This matches (qualitatively) the findings of Capra et al. [1999].[1]

2.1 Formal Definitions

Let $G = ([n], A, \vec{u})$ be a strategic game. We define iterated regret minimization in a way that makes it clear how it relates to other solution concepts based on iterated deletion. In particular, in the full paper [Halpern and Pass, Halpern and Pass2008], we relate it to rationalizability, and iterated deletion of weakly and strongly dominated strategies. A *deletion operator* $\mathcal{D}$ maps sets $\mathcal{S} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_n$ of strategy profiles in G to sets of strategy profiles such that $\mathcal{D}(\mathcal{S}) \subseteq \mathcal{S}$. We require that $\mathcal{D}(\mathcal{S}) = \mathcal{D}_1(\mathcal{S}) \times \cdots \times \mathcal{D}_n(\mathcal{S})$, where $\mathcal{D}_i$ maps sets of strategy profiles to strategies for player i. Intuitively, $\mathcal{D}_i(\mathcal{S})$ is the set of strategies for player i that survive deletion, starting with $\mathcal{S}$. Note that the set of strategies that survive deletion may depend on the set that we start with. Iterated deletion then amounts to applying the $\mathcal{D}$ operator repeatedly, starting with an appropriate initial set $\mathcal{S}_0$ of strategies, where $\mathcal{S}_0$ is typically either the set of pure strategy profiles (i.e., action profiles) in G or the set of mixed strategy profiles in G.

Definition 2.1 Given a deletion operator $\mathcal{D}$ and an initial set $\mathcal{S}_0$ of strategies, the set of strategy profiles that *survive iterated deletion with respect to $\mathcal{D}$ and $\mathcal{S}_0$* is

$$\mathcal{D}^\infty(\mathcal{S}_0) = \cap_{k>0} \mathcal{D}^k(\mathcal{S}_0)$$

(where $\mathcal{D}^1(\mathcal{S}) = \mathcal{D}(\mathcal{S})$ and $\mathcal{D}^{k+1}(\mathcal{S}) = \mathcal{D}(\mathcal{D}^k(\mathcal{S}))$. Similarly, the set of *strategy profiles for player i that survive iterated deletion with respect to $\mathcal{D}$ and $\mathcal{S}_0$* is $\mathcal{D}_i^\infty(\mathcal{S}_0) = \cap_{k>0} \mathcal{D}_i^k(\mathcal{S}_0)$, where $\mathcal{D}_i^1 = \mathcal{D}_i$ and $\mathcal{D}_i^{k+i} = \mathcal{D}_i \circ \mathcal{D}^k$.

---

[1]Capra et al. actually considered a slightly different game where the minimum bid was p (rather than 2). If we instead consider this game, we get an even closer qualitative match to their experimental observations.

We can now define the deletion operator $\mathcal{RM}$ appropriate for regret minimization in strategic games (we deal with Bayesian games in Section 3). Intuitively, $\mathcal{RM}_i(\mathcal{S})$ consists of all the strategies in $\mathcal{S}_i$ that minimize regret, given that the other players are using a strategy in $\mathcal{S}_{-i}$. In more detail, we proceed as follows. Suppose that G is a strategic game $([n], A, \vec{u})$ and that $\mathcal{S} \subseteq A$, the set of pure strategy profiles (i.e., actions). For $\vec{a}_{-i} \in \mathcal{S}_{-i}$, let $u_i^{\mathcal{S}_i}(\vec{a}_{-i}) = \max_{a_i \in \mathcal{S}_i} u_i(a_i, \vec{a}_{-i})$. Thus, $u_i^{\mathcal{S}_i}(\vec{a}_{-i})$ is the best outcome for i given that the remaining players play $\vec{a}_{-i}$ and that i can select actions only in $\mathcal{S}_i$. For $a_i \in \mathcal{S}_i$ and $\vec{a}_{-i} \in \mathcal{S}_{-i}$, let the *regret* of a_i for player i given $\vec{a}_{-i}$ relative to $\mathcal{S}_i$, denoted $regret_i^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$, be $u_i^{\mathcal{S}_i}(\vec{a}_{-i}) - u_i(a_i, \vec{a}_{-i})$. Let $regret_i^{\mathcal{S}}(a_i) = \max_{\vec{a}_{-i} \in \mathcal{S}_{-i}} regret^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$ denote the maximum regret of a_i for player i (given that the other players' actions are chosen from $\mathcal{S}_{-i}$). Let $minregret_i^{\mathcal{S}} = \min_{a_i \in \mathcal{S}_i} regret_i^{\mathcal{S}_{-i}}(a_i)$ be the minimum regret for player i relative to $\mathcal{S}$. Finally, let

$$\mathcal{RM}_i(\mathcal{S}) = \{a_i \in \mathcal{S}_i : regret_i^{\mathcal{S}}(a_i) = minregret_i^{\mathcal{S}}\}.$$

Thus, $\mathcal{RM}_i(\mathcal{S})$ consists of the set of actions that achieve the minimal regret with respect to $\mathcal{S}$. Clearly $\mathcal{RM}_i(\mathcal{S}) \subseteq \mathcal{S}$. Let $\mathcal{RM}(\mathcal{S}) = \mathcal{RM}_1(\mathcal{S}) \times \cdots \times \mathcal{RM}_n(\mathcal{S})$.

If $\mathcal{S}$ consists of mixed strategies, then the construction of $\mathcal{RM}(\mathcal{S})$ is the same, except that the expected utility operator U_i is used rather than u_i in defining $regret_i$.

Definition 2.2 Let $G = ([n], A, \vec{u})$ be a strategic game. $\mathcal{RM}_i^{\infty}(A)$ is the set of *(pure) strategies for player i that survive iterated regret minimization with respect to pure strategies in G*. Similarly, $\mathcal{RM}_i^{\infty}(\Sigma(A))$ is the set of *(mixed) strategies for player i that survive iterated regret minimization with respect to mixed strategies in G*.

The following theorem, proved in the full paper [Halpern and Pass, Halpern and Pass2008], shows that iterated regret minimization is a reasonable concept in that, for all games G, $\mathcal{RM}^{\infty}(A)$ and $\mathcal{RM}^{\infty}(\Sigma(A))$ are nonempty fixed points of the deletion process, that is, $\mathcal{RM}(\mathcal{RM}^{\infty}(A)) = \mathcal{RM}^{\infty}(A)$ and $\mathcal{RM}(\mathcal{RM}^{\infty}(\Sigma(A))) = \mathcal{RM}^{\infty}(\Sigma(A))$; the deletion process converges at $\mathcal{RM}^{\infty}$. (Our proof actually shows that for any nonempty closed set $\mathcal{S}$ of strategies, the set $\mathcal{RM}^{\infty}(\mathcal{S})$ is nonempty and is a fixed point of the deletion process.)

Theorem 2.3 *Let $G = ([n], A, \vec{u})$ be a strategic game. If $\mathcal{S}$ is a closed, nonempty set of strategies of the form $\mathcal{S}_1 \times \ldots \times \mathcal{S}_n$, then $\mathcal{RM}^{\infty}(\mathcal{S})$ is nonempty, $\mathcal{RM}^{\infty}(\mathcal{S}) = \mathcal{RM}_1^{\infty}(\mathcal{S}) \times \ldots \times \mathcal{RM}_n^{\infty}(\mathcal{S})$, and $\mathcal{RM}(\mathcal{RM}^{\infty}(\mathcal{S})) = \mathcal{RM}^{\infty}(\mathcal{S})$.*

Unlike solution concepts that implicitly assume that agents know other agents' strategies, in a strategy profile that survives iterated regret minimization, a player is not making a best response to the strategies used by the other players since, intuitively, he does not know what these strategies are. As a result, a player chooses a strategy that ensures that he does reasonably well compared to the best he could have done, no matter what the other players do. We shall see the impact of this in the examples of Section 2.3.

2.2 An Epistemic Characterization

Traditional solution concepts typically assume common knowledge of rationality, or at least a high degree of mutual knowledge of rationality. (This point is discussed in more detail in the full paper.) But if we make this assumption (and identify rationality with minimizing regret), we seem to run into a serious problem with Iterated Regret Minimization, which is well illustrated by the Traveler's Dilemma. As we observed earlier, the strategy profile $(97, 97)$ is the only one that survives iterated regret minimization when $p = 2$. However, if agent 1 knows that player 2 is playing 97, then he should play 96, not 97! That is, among all strategies, 97 is certainly not the strategy minimizes regret with respect to $\{97\}$.

In the full paper, we give an epistemic characterization of Iterated Regret Minimization that deals with this problem. We sketch the main ideas here.

Some of these difficulties also arise when dealing with iterated deletion of weakly dominated strategies. The justification for deleting a weakly dominated strategy is the existence of other strategies. But this justification may disappear in later deletions. As Mas-Colell, Whinston, and Green [1995, p. 240] put it:

> [T]he argument for deletion of a weakly dominated strategy for player i is that he contemplates the possibility that every strategy combination of his rivals occurs with positive probability. However, this hypothesis clashes with the logic of iterated deletion, which assumes, precisely, that eliminated strategies are not expected to occur.

Brandenburger, Friedenburg, and Kiesler [2008] resolve this paradox in the context of iterated deletion of weakly dominated strategies by assuming that strategies were not really eliminated. Rather, they assumed that strategies that are weakly dominated occur with infinitesimal (but nonzero) probability.

Unfortunately, this approach does not seem to help in the context of iterated regret minimization. Assigning deleted strategies infinitesimal probability will not make 97 a best response to a set of strategies where 97 is given very high probability. We deal with this problem by essentially reversing the approach taken by Brandenburger, Friedenberg, and Keisler. Rather than assuming common knowledge of rationality, we assign successively lower probability to higher orders of rationality. The idea is that now, with overwhelming probability, no assumptions are made about the other players; with probability ϵ, they are assumed to be rational, with probability ϵ^2, the other players are assumed to be rational and to believe that they are playing rational players, and so on. (Of course, "rationality" is interpreted here as minimizing expected regret.) Thus, for example, in Traveler's Dilemma, players do not play 96 because they are still assuming that, with overwhelming likelihood, the other player is playing an arbitrary strategy (not 97); 97 is slightly better than the other strategies that minimize regret given the slight likelihood that the other player is minimizing regret. As we show in the full paper, this approach does indeed justify iterated regret minimization.

155

2.3 Examples

We now consider the outcome of iterated regret minimization in a number of standard games, showing how it compares to the strategies recommended by other solution concepts. Due to lack of space, we focus here on just a few games and restrict to pure strategies; in the full paper, we consider a number of other examples, and also consider mixed strategies.

Example 2.4 *Traveler's Dilemma:* If $G = ([n], A, \vec{u})$ is the Traveler's Dilemma, then using the arguments sketched in the introduction, we get that $\mathcal{RM}_i^\infty(A) = \mathcal{RM}_i^2(A) = \{100 - 2p + 1\}$ if $p \geq 50$. As we mentioned, $(2, 2)$ is the only action profile (and also the only mixed strategy profile) that is rationalizable (resp., survives iterated deletion of weakly dominated strategies, is a Nash equilibrium). Moreover, if we allow an action to be deleted if it is strongly dominated by a mixed strategy, then $(2, 2)$ is the only action profile that survives iterated deletion of strongly dominated strategies. (This is not surprising, since it is well known [Osborne and Rubinstein, Osborne and Rubinstein1994] that a strategy survives iterated deletion of strongly dominated strategies iff it is rationalizable.) Thus, iterated regret minimization is quite different from all these other approaches in the Traveler's Dilemma, and gives results that are in line with experimental observations. ∎

Example 2.5 *Centipede Game:* Another well-known game for which traditional solution concepts provide an answer that is not consistent with empirical observations is the *Centipede Game* [Rosenthal, Rosenthal1982]. In the Centipede Game, two players play for a fixed number k of rounds (known at the outset). They move in turn; the first player moves in all odd-numbered rounds, while the second player moves in even-numbered rounds. At her move, a player can either stop the game, or continue playing (except at the very last step, when a player can only stop the game). For all t, player 1 prefers the stopping outcome in round $2t + 1$ (when she moves) to the stopping outcome in round $2t + 2$; similarly, for all t, player 2 prefers the outcome in round $2t$ (when he moves) to the outcome in round $2t + 1$. However, for all t, the outcome in round $t + 2$ is better for both players than the outcome in round t.

Consider two versions of the Centipede Game. The first has *exponential* payoffs. In this case, the utility of stopping at odd-numbered rounds t is $(2^t + 1, 2^{t-1})$, while the utility of stopping at even-numbered rounds is $(2^{t-1}, 2^t + 1)$. Thus, if player 1 stops at round 1, player 1 gets 3 and player 2 gets 1; if player 2 stops at round 4, then player 1 gets 8 and player 2 gets 17; if player 2 stops at round 20, the both players get over 500,000. In the version with *linear payoffs with punishment* $p > 1$, if t is odd, the payoff is $(t, t - p)$, while if t is even, the payoff is $(t - p, t)$.

The game can be described as a strategic game where A_i is the set of strategies for player i in the extensive-form game. It is straightforward to show (by backwards induction) that the only strategy profiles that survive iterated deletion of weakly dominated strategies are ones where player 1 stops at the first move and player 2 stops at the second move. Moreover, in all Nash equilibria, the first player stops at the first move. By way of contrast, it is not hard to show that all pure strategies where player 1 quits at some point (including the one where he continues until to his last move, and then quits) are rationalizable. Thus, Nash equilibrium seems to eliminate too many strategies, while rationalizabilitiy eliminates too few. In empirical tests (which have been done with linear payoffs), subjects usually cooperate for a certain number of rounds, although it is rare for them to cooperate throughout the whole game [McKelvey and Palfrey, McKelvey and Palfrey1992; Nagel and Tang, Nagel and Tang1998]. As we show in the full paper, using iterated regret minimization, with linear payoffs, we also get cooperation for a number of rounds (which depends on the penalty); with exponential payoffs, we get cooperation up to the end of the game. Our results suggest some further experimental work, with regard to the sensitivity of the game to the payoffs. ∎

Example 2.6 *Coordination games:* Suppose that $A_1 = A_2 = \{a, b\}$, and $u(a, a) = (k_1, k_2), u(b, b) = (1, 1), u(a, b) = u(b, a) = (0, 0)$, where $k_1, k_2 > 0$. Both (a, a) and (b, b) are Nash equilibria, but (a, a) *Pareto dominates* (b, b) if $k_1, k_2 > 1$: both players are better off with the equilibrium (a, a) than with (b, b). With regret minimization, we do not have to appeal to Pareto dominance if we stick to pure strategies. It is easy to see that if $k_1, k_2 > 1$, $\mathcal{RM}_1^1(A) = \mathcal{RM}_2^1(A) = \mathcal{RM}_1^\infty(A) = \mathcal{RM}_2^\infty(A) = \{a\}$ (yielding regret 1). Similarly, if $k_1, k_2 < 1$, then $\mathcal{RM}_1^\infty(A) = \mathcal{RM}_2^\infty(A) = \{b\}$, and if if $k_1 = k_2 = 1$, $\mathcal{RM}_1^1(A) = \mathcal{RM}_2^1(A) = \mathcal{RM}_1^\infty(A) = \mathcal{RM}_2^\infty(A) = \{a, b\}$. Finally, if $k_1 > 1$, and $k_2 < 1$, then the unique profile that minimizes regret is (a, b), which results in both players getting a payoff of 0. While this is unfortunate, in the absence of any coordinating device, this may well be the outcome that is played between two players meeting for the first time. ∎

3 Bayesian Games

Bayesian games are a well-known generalization of strategic games, where each agent is assumed to have a characteristic or some private information not known to the other players. This is modeled by assuming each player has a *type*. Typically it is assumed that that there is a commonly known probability over the set of possible type profiles. Thus, a Bayesian game is tuple $([n], A, \vec{u}, T, \pi)$, where, as before, $[n]$ is the set of players, A is the set of action profiles, $\vec{u}$ is the profile of utility functions, $T = T_1 \times \ldots \times T_n$ is the set of type profiles (where T_i represents the set of possible types for player i), and π is a probability measure on T. A player's utility can depend, not just on the action profile, but on the type profile. Thus, $u_i : A \times T \to I\!R$. For simplicity, we assume that $\Pr(t_i) > 0$ for all types $t_i \in T_i$ and $i = 1, \ldots, n$ (where t_i is an abbreviation of $\{\vec{t}' : t_i' = t_i\}$).

A strategy for player i in a Bayesian game in a function from player i's type to an action in A_i; that is, what a player does will in general depends on his type. For a pure strategy profile $\vec{\sigma}$, player i's expected utility is

$$U_i(\vec{\sigma}) = \sum_{\vec{t} \in T} \pi(\vec{t}) u_i(\sigma_1(t_1), \ldots, \sigma_n(t_n)).$$

Player i's expected utility with a mixed strategy profile $\vec{\sigma}$ is computed by computing the expectation with respect to the probability on pure strategy profiles induced by $\vec{\sigma}$. Given these definitions, a Nash equilibrium in a Bayesian game is defined in the same way as a Nash equilibrium in a strategic game.

There are some subtleties involved in doing iterated deletion in Bayesian games. Roughly speaking, we need to relativize all the previous definitions so that they take types into account. We give the definitions for pure strategies; the modifications to deal with mixed strategies are straightforward and left to the reader.

As before, suppose that $\mathcal{S} = \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$. Moreover, suppose that, for each player i, $\mathcal{S}_i$ is also a crossproduct; that is, for each type $t \in T_i$, there exists a set of actions $A(t) \subseteq A_i$ such that $\mathcal{S}_i$ consists of all strategies σ such that $\sigma(t) \in A(t)$ for all $t \in T_i$. For $\vec{a}_{-i} \in \mathcal{S}_{-i}$ and $\vec{t} \in T$, let $u_i^{\mathcal{S}_i}(\vec{a}_{-i}, \vec{t}) = \max_{a_i \in \mathcal{S}} u_i(a_i, \vec{a}_{-i}, \vec{t})$. For $a_i \in \mathcal{S}_i$, $\vec{a}_{-i} \in \mathcal{S}_{-i}$, and $\vec{t} \in T$, the *regret* of a_i for player i given $\vec{a}_{-i}$ and $\vec{t}$, relative to $\mathcal{S}_i$, denoted $regret_i^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i}, \vec{t})$, is $u_i^{\mathcal{S}_i}(\vec{a}_{-i}, \vec{t}) - u_i(a_i, \vec{a}_{-i}, \vec{t})$. Let $regret_i^{\mathcal{S}}(a_i \mid \vec{t}) = \max_{\vec{a}_{-i} \in \mathcal{S}_{-i}(\vec{t}_{-i})} regret^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i}, \vec{t})$ denote the maximum regret of player i given $\vec{t}$. The *expected regret of a_i given t_i and $\mathcal{S}_{-i}$* is $E[regret_i^{\mathcal{S}_i}(a_i \mid t_i)] = \sum_{\vec{t} \in T} \Pr(\vec{t} \mid t_i) regret_i^{\mathcal{S}_i}(a_i \mid \vec{t})$. Let $minregret^{\mathcal{S}_i}(t_i) = \min_{a_i \in \mathcal{S}_i(t_i)} E[regret_i^{\mathcal{S}}(a_i \mid t_i)]$. We delete all those strategies that do not give an action that minimizes expected regret for each type. Thus, let $\mathcal{RM}_i(\mathcal{S}_i) = \{\sigma \in \mathcal{S}_i : regret_i^{\mathcal{S}_i}(\sigma(t_i) \mid t_i) = minregret^{\mathcal{S}_i}(t_i)\}$, and define $\mathcal{RM}(\mathcal{S}) = \mathcal{RM}_1(\mathcal{S}_1) \times \ldots \times \mathcal{RM}_n(\mathcal{S}_n)$. We can then apply iterated deletion to this deletion operator.

Example 3.1 *Second-Price Auction:* A second-price auction can be modeled as a Bayesian game, where a player's type is his valuation of the product being auctioned. His possible actions are bids. The player with the highest bid wins the auction, but pays only what the second-highest (For simplicity, we assume that in the event of a tie, the lower-numbered player wins the auction.) If he bids b and has valuation (type) v, his utility is $v - b$; if he does not win the auction, his utility is 0. As is well known, in a second-price auction, the strategy where each player bids his type is weakly dominant; hence, this strategy survives iterated regret minimization. No other strategy can give a higher payoff, no matter what the type profile is. ∎

Example 3.2 *First-Price Auction:* In a first-price auction, the player with the highest bid wins the auction, but pays his actual bid. Assume, for simplicity, that bids are natural numbers, that the lower-numbered player wins the auction in the event of a tie, that all valuations are even, and that the product is sold only if some player bids above 0. If a player's valuation is v, then bidding v' has regret $max(v' - 1, v - v' - 1)$. To see this, consider player i. Suppose that the highest bid of the other agents is v'', and that the highest-numbered agent that bids v'' is agent j. If $v'' < v'$ or $v'' = v'$ and $i < j$, then i wins the bid. He may have done better by bidding lower, but the lowest he can bid and still win is 1, so his maximum regret in this case is $v' - 1$ (which occurs, for example, if $v'' = 0$).

On the other hand, if $v'' > v'$ or $v'' = v$ and $j < i$, then i does not win the auction. He feels regret if he could have won the auction and still paid at most v. To win, if $j < i$, he must bid $v'' + 1$, in which case his regret is $v - v'' - 1$. In this case, v'' can be as small as v', so his regret is at most $v - v' - 1$. If $j > i$, then he must only bid v'' to win, so his regret is $v - v''$, but $v'' \geq v' + 1$, so his regret is again at most $v - v' - 1$. It follows that bidding $v' = v/2$ is the unique action that minimizes regret (yielding a regret of $v/2 - 1$). ∎

In the full paper, we use regret minimization to analyze revenue-maximizing mechanisms for combinatorial auctions.

4 Related Work

While the notion of regret has been well studied in the context of decision theory (see [Hayashi, Hayashi2008] and the references therein for some discussion of the recent work), there has been surprisingly little work on applying regret to game theory. Linhart and Radner [1989] applied regret to a bargaining problem, and showed that it leads to solutions that seem much more reasonable than Nash equilibrium. Hyafil and Boutilier [2004] consider *pre-Bayesian* games, where each agent has a type and a player's utility depends on both the action profile and the type profile, just as in a Bayesian game, but now there is no probability on types.[2] The solution concept they use is a hybrid of Nash equilibrium and regret. Roughly speaking, they take regret with respect to the types of the other players, but use Nash equilibrium with respect to the strategies of other players. That is, they define $regret_i^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i}, \vec{t})$ as we do (taking $\mathcal{S}_i$ to consist of all strategies for player i), but then define $regret^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$ by minimizing over all $\vec{t}_{-i}$. They then define a profile $\vec{\sigma}$ to be a *minimax-regret equilibrium* if, for all type profiles $\vec{t}$, no agent can decrease his regret by changing his action. For strategic games, where there are no types (i.e., $|T| = 1$), their solution concept collapses to Nash equilibrium. Thus, their definitions differ from ours in that they take regret with respect to types, not with respect to the strategies of other players as we do, and they do not iterate the regret operation.

Even closer to our work is a recent paper by Renou and Schlag [2008] (done independently of ours). Just as we do, they focus on strategic games. Although they define regret for pure strategies, this is only a tool for dealing with mixed strategies; they do not consider the regret of a pure strategy with respect to a set of pure strategies, as we do (and they do not consider iteration). In particular, they have no analogue to our analysis of pure strategies in Section 2.3. Moreover, their definition of regret for mixed strategies is somewhat different from ours; see the full paper for a detailed comparison.

5 Conclusion

There are many examples in the literature of games where Nash equilibrium and its refinements do not describe what

---

[2]Hyafil and Boutilier actually consider a slightly less general setting, where the utility for player i depends only on player i's type, not the whole type profile. Modifying their definitions for the more general setting is straightforward.

people do. We have introduced a new solution concept, iterated regret minimization, that, at least in some games, seems to capture better what people are doing than more standard solution concepts. The outcomes of games like the Traveler's Dilemma and the Centipede Game have sometimes been explained by assuming that a certain fraction of agents will be "altruistic", and play the helpful action (e.g., playing 100 in Traveler's Dilemma or cooperating in the Centipede Game) (cf., [Capra, Goeree, Gomez, and Holt, Capra et al.1999]). Altruism may indeed be part of an accurate descriptive theory, but to use it, we first need to decide what the "right" action is, and also with what percentage agents are altruistic. We also need to explain why this percentage may depend on the degree of punishment (as it did in the experiments of Capra et al. [1999], for example). Iterated regret minimization provides a different descriptive explanation, and has some normative import as well. It seems particularly appealing when considering inexperienced but intelligent players that play a game for the first time. In this setting, it seems unreasonable to assume that players know what strategies other players are using (as is implicitly assumed in Nash equilibrium).

References

Aumann, R. J. and A. Brandenburger (1995). Epistemic conditions for Nash equilibrium. *Econometrica 63*(5), 1161–1180.

Basu, K. (1994). The traveler's dilemma: paradoxes of rationality in game theory. *American Economic Review 84*(2), 391–395.

Basu, K. (2007). The traveler's dilemma. *Scientific American* June, 90–95.

Becker, T., M. Carter, and J. Naeve (2005). Experts playing the Traveler's Dilemma. Discussion paper 252/2005, Universität Hohenheim.

Brandenburger, A., A. Friedenberg, and J. Keisler (2008). Admissibility in games. *Econometrica 76*(2), 307–352.

Capra, M., J. K. Goeree, R. Gomez, and C. A. Holt (1999). Anamolous behavior in a traveler's dilemma. *American Economic Review 89*(3), 678–690.

Halpern, J. Y. and R. Pass (2008). Iterated regret minimization: A more realistic solution concept. arxiv.org/abs/0810.3023.

Hayashi, T. (2008). Regret aversion and opportunity dependence. *J. Economic Theory 139*(1), 242–268.

Hyafil, N. and C. Boutilier (2004). Regret minimizing equilibria and mechanisms for games with strict type uncertainty. In *Proc. 20th Conf. Uncertainty in Artificial Intelligence (UAI 2004)*, pp. 268–277.

Linhart, P. B. and R. Radner (1989). Minimax-regret strategies for bargaining over several variable. *J. Economic Theory 48*, 152–178.

Mas-Colell, A., M. Whinston, and J. Green (1995). *Microeconomic Theory*. Oxford University Press.

McKelvey, R. and T. Palfrey (1992). An experimental study of the centipede game. *Econometrica 60*(4), 803–836.

Nagel, R. and F. F. Tang (1998). An experimental study on the centipede game in normal form—An investigation on learning. *J. Mathematical Psychology 42*, 239–265.

Niehans, J. (1948). Zur preisbildung bei ungewissen erwartungen. *Scbweizerische Zietschrift für Volkswirtschaft und Statistik 84*(5), 433–456.

Osborne, M. J. and A. Rubinstein (1994). *A Course in Game Theory*. MIT Press.

Renou, L. and K. H. Schlag (2008). Minimax regret and strategic uncertainty. Working Paper No. 08/2, University of Leicester, Department of Economics.

Rosenthal, R. W. (1982). Games of perfect information, predatory pricing, and the chain store paradox. *J. Economic Theory 25*, 92–100.

Savage, L. J. (1951). The theory of statistical decision. *J. American Statistical Association 46*, 55–67.