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. Other solution concepts have been considered, such as *sequential equilibrium* [Kreps and Wilson 1982], *perfect equilibrium* [Selten 1975], and *rationalizability* [Bernheim 1984; Pearce 1984]; see [Osborne and Rubinstein 1994] 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 1994; Basu 2007]. 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.

The rest of this paper is organized as follows. Section 2 contains preliminaries. Section 3 is the heart of the paper: we first define iterated regret minimization in strategic games, provide an epistemic characterization of it, and then show how iterated regret minimization works in numerous standard examples from the game theory literature, including Traveler's Dilemma, Prisoner's Dilemma, the Centipede Game, Bertrand Competition, and Nash Bargaining, both with pure strategies and mixed strategies. The epistemic characterization, like those of many other solution concepts, involves higher and higher levels of belief regarding other players' rationality, it does *not* involve common knowledge or common belief. Rather, higher levels of beliefs are accorded lower levels of likelihood. In Sections 4 and 5, we briefly consider regret minimization in Bayesian games and in the context of mechanism design. We discuss related work in Section 6, and conclude in Section 7. Proofs are relegated to the appendix.

2 Preliminaries

We refer to a collection of values, one for each player, as a *profile*. If player j's value is x_j, then the resulting profile is denoted $(x_j)_{j\in[n]}$, or simply (x_j) or $\vec{x}$, if the set of players is clear from the context. Given a profile $\vec{x}$, let $\vec{x}_{-i}$ denote the collection consisting of all values x_j for $j \neq i$. It is sometimes convenient to denote the profile $\vec{x}$ as $(x_i, \vec{x}_{-i})$.

A *strategic game in normal form* is a "single-shot" game, where each player i chooses an action from a space A_i of actions. For simplicity, we restrict our attention to *finite* games—i.e., games where the set A_i is finite. Let $A = A_1 \times \ldots \times A_n$ be the set of action profiles. A strategic game is characterized by a tuple $([n], A, \vec{u})$, where $[n]$ is the set of players, A is the set of action profiles, and $\vec{u}$ is the profile of utility functions, where $u_i(\vec{a})$ is player i's utility or payoff if the action profile $\vec{a}$ is played. A *(mixed) strategy* for player i is a probability distribution $\sigma_i \in \Delta(A_i)$ (where, as usual, we denote by $\Delta(X)$ the set of distributions on the set X). Let $\Sigma_i = \Delta(A_i)$ denote the mixed strategies for player i in game G, and let $\Sigma = \Sigma_1 \times \cdots \times \Sigma_n$ denote the set of mixed strategy profiles. Note that, in strategy profiles in Σ, players are randomizing independently.

A *pure strategy* for player i is a strategy for i that assigns probability 1 to a single action. To simplify notation, we let an action $a_i \in A_i$ also denote the pure strategy $\sigma_i \in \Delta(A_i)$ which puts weight only on a_i. If σ is a strategy for player i then $\sigma(a)$ denotes the probability given to action a by strategy σ. Given a strategy profile $\vec{\sigma}$, player i's expected utility if $\vec{\sigma}$ is played, denoted $U_i(\vec{\sigma})$, is $\mathbf{E}_{\Pr}[u_i^{\vec{\sigma}}]$, where the expectation is taken with respect to the probability Pr induced by $\vec{\sigma}$ (where the players are assumed to choose their actions independently).

3 Iterated Regret Minimization in Strategic Games

We start by providing an informal discussion of iterated regret minimization in strategic 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. 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 $\mathit{regret}_u(a, s)$, 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 $\mathit{regret}_u(a) = \max_{s \in S} \mathit{regret}_u(a, s)$. For example, if $\mathit{regret}_u(a) = 2$, 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 other work on applying regret minimization to game theory in Section 6; 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 $2p-1$. In the special case that $p=2$, 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, at a qualitative level, the findings of Capra et al. [1999].[1]

3.1 Deletion Operators and Iterated Regret Minimization

Iterated regret minimization proceeds much like other notions of iterated deletion. To put it in context, we first abstract the notion of iterated deletion.

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. 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 3.1 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

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

[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.

(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 $\mathcal{D}_i^{k+i} = \mathcal{D}_i \circ \mathcal{D}^k$.

We can now define the deletion operator $\mathcal{RM}$ appropriate for regret minimization in strategic games (we deal with Bayesian games in Section 4). 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 $\mathit{regret}_i^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$, be $u_i^{\mathcal{S}_i}(\vec{a}_{-i}) - u_i(a_i, \vec{a}_{-i})$. Let $\mathit{regret}_i^{\mathcal{S}}(a_i) = \max_{\vec{a}_{-i} \in \mathcal{S}_{-i}} \mathit{regret}^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$ denote the maximum regret of a_i for player i (given that the other players' actions are chosen from $\mathcal{S}_{-i}$). Let $\mathit{minregret}_i^{\mathcal{S}} = \min_{a_i \in \mathcal{S}_i} \mathit{regret}_i^{\mathcal{S}_{-i}}(a_i)$ be the minimum regret for player i relative to $\mathcal{S}$. Finally, let

$$\mathcal{RM}_i(\mathcal{S}) = \{a_i \in \mathcal{S}_i : \mathit{regret}_i^{\mathcal{S}}(a_i) = \mathit{minregret}_i^{\mathcal{S}}\}.$$

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. We also need to argue that there is a strategy s_i for player i that maximizes regret_i and one that minimizes $\mathit{minregret}_i$. This follows from the compactness of the sets of which the max and min are taken, and the continuity of the functions being maximized and minimized.

Definition 3.2 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, whose proof is in the appendix, 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 3.3 *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 3.4.

3.2 Comparison to Other Solution Concepts Involving Iterated Deletion

Iterated deletion has been applied in other solution concepts. We mention three here. Given a set $\mathcal{S}$ of strategies, a strategy $\sigma \in \mathcal{S}_i$ is *weakly dominated* by $\tau \in \mathcal{S}_i$ with respect to $\mathcal{S}$ if, for some strategy $\vec{\sigma}_{-i} \in \mathcal{S}_{-i}$, we have $U_i(\sigma, \vec{\sigma}_{-i}) < U_i(\tau, \vec{\sigma}_{-i})$ and, for all strategies $\vec{\sigma}'_{-i} \in \mathcal{S}_{-i}$, we have $U_i(\sigma, \vec{\sigma}'_{-i}) \leq U_i(\tau, \vec{\sigma}'_{-i})$. Similarly, σ is *strongly dominated* by τ with respect to $\mathcal{S}$ if $U_i(\sigma, \vec{\sigma}'_{-i}) < U_i(\tau, \vec{\sigma}'_{-i})$ for all strategies $\vec{\sigma}'_{-i} \in \mathcal{S}_{-i}$. Thus, if σ is weakly dominated by τ with respect to $\mathcal{S}_{-i}$, then i always does at least as well with τ as with σ, and sometimes does better (given that we restrict to strategies in $\mathcal{S}_{-i}$); if σ is strongly dominated by τ, then player i always does better with τ as with σ. Let $\mathcal{WD}_i(\mathcal{S})$ (resp., $\mathcal{SD}_i(\mathcal{S})$) consist of all strategies $\sigma_i \in \mathcal{S}_i$ that are not weakly (resp., strongly) dominated by some strategy in $\mathcal{S}_i$ with respect to $\mathcal{S}$. We can then define the *pure strategies that survive iterated weak (resp., strong) deletion with respect to pure strategies* as $\mathcal{WD}(A)$ (resp., $\mathcal{SD}(A)$). And again, we can start with Σ to get corresponding notions for mixed strategies.

As is well known [Osborne and Rubinstein 1994], the *rationalizable* strategies can also be considered as the outcome of an iterated deletion process. Intuitively, a pure strategy for player i is rationalizable if it is a best response to some beliefs that player i may have about the pure strategies that other players are following. Given a set $\mathcal{S}$ of pure strategy profiles, $\sigma \in \mathcal{S}_i$ is *justifiable* if there is some distribution μ on the strategies in $\mathcal{S}_{-i}$ such that σ is a best response to the resulting mixed strategy. Intuitively, μ describes player i's beliefs about the likelihood that other players are following various strategies; thus, a strategy σ for i is justifiable if there are beliefs that i could have to which σ is a best response. Let $\mathcal{J}_i(\mathcal{S})$ consist of all strategies for player i that are justifiable with respect to $\mathcal{S}$. A pure strategy σ for player i is *rationalizable* if $\sigma \in \mathcal{J}_i^\infty(A)$.[2]

3.3 An Epistemic Characterization of Iterated Regret Minimization

Traditional solution concepts typically assume common knowledge of rationality, or at least a high degree of mutual knowledge of rationality. For example, it is well

[2]The notion of rationalizability is typically applied to pure strategies, although the definitions can be easily extended to deal with mixed strategies.

known that rationalizability can be characterized in terms of common knowledge of rationality [Tan and Werlang 1988], where rational if he has some beliefs according to which what he does is a best response in terms of maximizing expected utility; Aumann and Brandenburger [1995] show that Nash equilibrium requires (among other things) mutual knowledge of rationality (where, again, rationality means playing a utility-maximizing best response); and Brandenburger, Friedenberg, and Kiesler [2008] show that iterated deletion of weakly dominated strategies requires sufficiently high mutual assumption of rationality, where "assumption" is a variant of "knowledge", and "rationality" means "does not play a weakly dominated strategy". 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 $\{97\}$.

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 in their textbook when discussing iterated deletion of weakly dominated strategies:

> [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. This is formally modeled in a framework where uncertainty is captured using a lexicographic probability system (LPS) [Blume, Brandenburger, and Dekel 1991], whose support consists of all types. (Recall that an LPS is a sequence $(\mu_0, \mu_1, \ldots)$ of probability measures, in this case on type profiles, where μ_1 represents events that have infinitesimal probability relative to μ_0, μ_1 represents events that have infinitesimal probability relative to μ_1, and so on. Thus, a probability of $(1/2, 1/3, 1/4)$ can be identified with a nonstandard probability of $1/2 + \epsilon/3 + \epsilon^2/4$, where ϵ is an infinitesimal.) In this framework, they show that iterated deletion of weakly dominated strategies corresponds to sufficiently high mutual assumption of rationality, where "assumption" is a variant of "knowledge", and "rationality" means "does not play a weakly dominated strategy".

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.

This approach is consistent with the spirit of Camerer, Ho, and Chong's [2004] *cognitive hierarchy* model, where the fraction of people with kth-order beliefs declines as a function of k, although not as quickly as this informal discussion suggests.

Since regret minimization is non-probabilistic, the formal model of a lexicographic belief is a countable sequence $(\mathcal{S}^0, \mathcal{S}^1, \ldots)$ of sets of strategy profiles. The strategy profiles in $\mathcal{S}^0$ represent the players' primary beliefs, the strategy profiles in $\mathcal{S}^1$ are the players' secondary beliefs, and so on. (We can think of $\mathcal{S}^k$ as the support of the measure μ_k in an LPS.)[3] We call $\mathcal{S}^i$ the *level-i belief* of the lexicographic belief $(\mathcal{S}^0, \mathcal{S}^1, \ldots)$.

Given such lexicographic beliefs, what strategy should a rational player i choose? Clearly the most important thing is to minimize regret with respect to his primary beliefs, $\mathcal{S}^0_{-i}$. But among strategies that minimize regret with respect to $\mathcal{S}^0_{-i}$, the best are those strategies that also minimize regret with respect to $\mathcal{S}^1_{-i}$; similarly, among strategies that minimize regret with respect to each of $\mathcal{S}^1_{-i}, \ldots, \mathcal{S}^{k-1}_{-i}$, the best are those that also minimize regret with respect to $\mathcal{S}^k_{-i}$. Formally, a strategy σ for player i is *rational with respect to a lexicographic sequence* $(\mathcal{S}^0, \mathcal{S}^1, \ldots)$ if there exists a sequence $(\mathcal{T}^0, \mathcal{T}^1, \ldots)$ of strategy profiles such that $\mathcal{T}^0_i$ consists of all strategies τ such that τ_i minimizes regret with respect to $\mathcal{S}^0_{-i}$ for all players i; and $\mathcal{T}^k$ for $k > 0$ is defined inductively to consist of all strategies $\tau \in \mathcal{T}^{k-1}$ such that τ_i has the least regret with respect to $\mathcal{S}^k_{-i}$ among all strategies in $\mathcal{T}^{k-1}_i$; and $\sigma \in \cap_{k=0}^{\infty} \mathcal{T}^k_i$.[4] Of course, this definition makes perfect sense if the lexicographic sequence is finite and has the form $(\mathcal{S}^0, \ldots, \mathcal{S}^k)$; in that case we consider $(\mathcal{T}^0, \ldots, \mathcal{T}^k)$. Such a sequence $(\mathcal{S}^0, \ldots, \mathcal{S}^k)$ is called a $(k+1)$*st-order lexicographic belief*. It easily follows that $\emptyset \neq \cdots \subseteq \mathcal{T}^k \subseteq \mathcal{T}^{k-1} \subseteq \cdots \subseteq \mathcal{T}^0$, so that a strategy that is rational with respect to $(\mathcal{S}^0, \mathcal{S}^1, \ldots)$ is also rational with respect to each of the finite prefixes $(\mathcal{S}^0, \mathcal{S}^1, \ldots)$.

Up to now, we have not imposed any constraints on *justifiability* of beliefs. We

[3]Like LPS's, this model implicitly assumes, among other things, that players i and j have the same beliefs about players $j' \notin \{i, j\}$. This assumption is acceptable, given that we assume that (it is commonly known that) all players start the iteration process by considering all strategies. To get an epistemic characterization of a more general setting, where players' initial beliefs about other players strategies are not commonly known, we need a slightly more general model of beliefs, where each player has his or her own lexicographic sequence; see Section 3.5.

[4]A straightforward argument by induction shows that $\mathcal{T}^k$ is nonempty and compact, so that there will be a strategy τ_i that has the least regret with respect to $\mathcal{S}^k_{-i}$ among all strategies in $\mathcal{T}^{k-1}_i$.

provide a recursive definition of justifiability. A (kth-order) lexicographic belief $(\mathcal{S}^j)_{j\in I}$ is *justifiable* if, for each $j \in I$, the level-j belief $\mathcal{S}_j$ is *level-j justifiable*, where level-j justifiability is defined as follows.

- To capture the intuition that players' primary beliefs are such that they make no assumptions about the other players, we say that a belief $\mathcal{S}_i^0$ is *level-0 justifiable* if it is the full set of strategies $\mathcal{S}_i$ available to player i.[5]

- To capture the intuition that players' level-k belief is that the other players are $(k-1)$st-order rational, we say that a belief, $\mathcal{S}_i^k$, is *level-k justifiable* if there exists some justifiable kth-order belief $(\mathcal{S}'^0, \mathcal{S}'^1, \ldots, \mathcal{S}'^{k-1})$ such that $\mathcal{S}_i^k$ is the set of rational strategies for player i with respect to $(\mathcal{S}'^0_{-i}, \mathcal{S}'^1_{-i}, \ldots, \mathcal{S}'^{k-1}_{-i})$.

This notion of justifiability captures the intuition that, with probability ϵ^k, each player j_k believes that each other player j_{k-1} is rational with respect to a kth-order belief and believes that, with probability ϵ^{k-1}, each other player j_{k-2} is rational with respect to a $(k-1)$st-order belief and believes that, with probability ϵ^{k-2}, ..., and with probability ϵ believes that each other player j_1 is rational with respect to a first-order belief and believes that each other player j_0 is playing an arbitrary strategy in $\mathcal{S}^0$ (As usual, "rationality" here means "minimizes regret with respect to his beliefs".)

Given these definition, we have the following theorem.

Theorem 3.4 *Let $G = ([n], A, \vec{u})$ be a strategic game and let $\mathcal{S}$ be the full set of pure or mixed strategies. Then for each $k \in N$ there exists a unique level-k justifiable belief $\tilde{\mathcal{S}}^k = \mathcal{RM}^{k-1}(\mathcal{S})$. Furthermore, $\mathcal{RM}^\infty(\mathcal{S})$ is the set of rational strategies with respect to the belief $(\tilde{\mathcal{S}}^0, \tilde{\mathcal{S}}^1, \ldots)$ and $\mathcal{RM}^k(\mathcal{S})$ is the set of rational strategies with respect to the belief $(\tilde{\mathcal{S}}^0, \tilde{\mathcal{S}}^1, \ldots, \tilde{\mathcal{S}}^k)$*

Proof: By definition there is a unique level-0 justifiable belief $\tilde{\mathcal{S}}^0 = \mathcal{S}$. It inductively follows that there exists a unique level-k justifiable belief $\tilde{\mathcal{S}}^k = \mathcal{RM}(\mathcal{S}^{k-1}) = \mathcal{RM}^{k-1}(\mathcal{S})$. The theorem then follows from the definition of rationality with respect to a lexicographic belief. ∎

Note that the sets $(\tilde{\mathcal{S}}^1, \tilde{\mathcal{S}}^2, \ldots)$ are just the sets $(\mathcal{T}^0, \mathcal{T}^1, \ldots)$ given in the definition of rationality.

In Appendix A, we provide an alternative characterization of iterated regret minimization in terms of Kripke structures. The idea is to say that, at level 0, each player has no idea what strategy the other players are using; at level 1, the player knows that the other players are rational, but that is all he knows; at level 2, the player knows that other players are rational and believe, at level 1, that other players are rational, but that is all he knows; and so on. This is made precise using ideas from [Halpern and Pass 2009].

[5] As we discuss in Section 3.5, we can also consider a more general model where players have prior beliefs; in such a setting, $\mathcal{S}_i^0$ need not be the full set of strategies.

3.4 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. We start by considering what happens if we restrict to pure strategies, and then consider mixed strategies.

3.4.1 Pure strategies

Example 3.5 *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 1994] 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.

Interestingly, although (2,2) is the only Nash equilibrium, it is a rather fragile equilibrium. If, rather than assuming common belief in rationality, we assume only common p-belief [Monderer and Samet 1989] for even a relatively small p such as .02 (where common p-belief of a statement ϕ holds if everyone believes with probability $1-p$ that everyone believes with probability $1-p$ ... that ϕ holds), then almost every bid can be justified. But this leaves us with the problem of choosing what will be played; iterated regret minimization provides some insight. ∎

Example 3.6 *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 1982]. 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, a large number of strategies are rationalizable. Indeed, 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. For example, it is rationalizable for player 2 to continue until the end (since he can believe that, with probability 1, player 1 will stop at the first move, so that his strategy is irrelevant). Given that this is rationalizabile, it is also rationalizable for player 1 to contine to his last move and then quit. 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 1992; Nagel and Tang 1998]. As we now show, 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.

Before going on, note that, technically, a strategy in the Centipede Game must specify what a player does whenever he is called upon to move, including cases where he is called upon to move after he has already stopped the game. Thus, if t is odd and $t + 2 < k$, there is more than one strategy where player 1 stops at round t. For example, there is one where player 1 also stops at round $t + 2$, and another where he continues at round $t+2$. However, all the strategies where player 1 first stops at round t are payoff equivalent for player 1 (and, in particular, are equivalent with respect to regret minimization). We use $[t]$ denote the set of strategies where player 1 stops at round t, and similarly for player 2. It is easy to see that in the k-round Centipede Game with exponential payoffs, the unique strategy that minimizes regret is to stop at the last possible round. On the other hand, with linear payoffs and punishment p, the situation is somewhat similar to the Traveler's Dilemma. All strategies (actions) for player 1 that stop at or after stage $k - p + 1$ have regret $p - 1$, which is minimal, but what happens with iterated deletion depends on whether k and p are even or odd. For example, if k and p are both even, then $\mathcal{RM}_1(A) = \{[k-p+1], [k-p+3], \ldots, k-1\}$ and $\mathcal{RM}_2(A) = \{k-p+2, k-p+4, \ldots, k\}$. Relative to $\mathcal{RM}(A)$, the strategies in $[k-p+1]$ have regret $p-2$; the remaining strategies have regret $p-1$. Thus, $\mathcal{RM}_1^2(A) = \mathcal{RM}_2^\infty(A) = \{[k-p+1]\}$. Similarly, $\mathcal{RM}_2^2(A) = \mathcal{RM}_2^\infty(A) = \{[k-p+2]\}$.

If, on the other hand, both k and p are odd, $\mathcal{RM}_1(A) = \{[k-p+1], [k-p+3], \ldots, [k]\}$ and $\mathcal{RM}_2(A) = \{[k-p], [k-p+2], \ldots, [k-1]\}$. But, here, iteration does not remove any strategies (as here $k-p+1$ still has regret $p-1$ for player 1, and $k-p$ still has regret $p-1$ for player 2. Thus, $\mathcal{RM}_1^\infty(A) = \mathcal{RM}_1(A)$ and $\mathcal{RM}_2^\infty(A) = \mathcal{RM}_2(A)$. ∎

Example 3.7 *Matching pennies:* Suppose that $A_1 = A_2 = \{a, b\}$, and $u(a, a) = u(b, b) = (80, 40), u(a, b) = u(b, a) = (40, 80)$, and consider the matching pennies game, with payoffs as given in the table below:

	a	b
a	$(80, 40)$	$(40, 80)$
b	$(40, 80)$	$(80, 40)$

Since the players have opposite interest there are no pure strategy Nash equilibria. Randomizing with equal probability over both actions is the only Nash equilibria; this is consistent with experimental results (see e.g., [Goeree and Holt 2001]). With regret minimization, both actions have identical regret for both players, thus using regret minimization (with respect to pure strategies) both actions are viable.

Consider a variant of this game (called the *asymmetric matching pennies* [Goeree, Holt, and Palfrey 2000]), where $u(a, a) = (320, 40)$. Here, the unique Nash equilibrium

	a	b
a	$(320, 40)$	$(40, 80)$
b	$(40, 80)$	$(80, 40)$

is one where player 1 still randomizes between a and b with equal probability, but player 2 picks b with probability 0.875 [Goeree, Holt, and Palfrey 2000]. Experimental results by Goeree and Holt [2001] show quite different results: player 1 chooses a with probability 0.96; player 2, on the other hand, is consistent with the Nash equilibrium and chooses b with probability 0.86. In other words, players most of the time end up with the outcome (a, b). With iterated regret minimization, we get a qualitatively similar result. It is easy to see that in the first round of deletion, a minimizes the regret for player 1, whereas both a and b minimize the regret for player 2; thus $\mathcal{RM}_1^1(A) = a$ and $\mathcal{RM}_2^1(A) = a, b$. In the second round of the iteration, b is the only action that minimize regret for player 2. Thus, $\mathcal{RM}^2(A) = \mathcal{RM}^\infty = (a, b)$; that is, (a, b) is the only strategy profile that survives iterated deletion. ∎

Example 3.8 *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 $k_1, k_2 > 0$, as shown in the table below:

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

	a	b
a	(k_1, k_2)	$(0,0)$
b	$(0,0)$	$(1,1)$

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. ▮

Example 3.9 *Bertrand competition:* Bertrand competition is a 2-player game where the players can be viewed as firms producing a homogeneous good. There is demand for 100 units of the good at any price up to \$200. If both firms charge the same price, each will sell 50 units of the good. Otherwise, the firm that charges the lower price will sell 100 units at that price. Each firm has a cost of production of 0, as long as it sells for a positive price, it makes a profit. It is easy to see that the only Nash equilibria of this game are (0,0) and (1,1). But it does not seem so reasonable that firms playing this game only once will charge \$1, when they could charge up to \$200. And indeed, experimental evidence shows that people in general choose significantly higher prices [Dufwenberg and Gneezy 2000].

Now consider regret. Suppose that firm 1 charges $n \geq 1$. If firm 2 charges $m > 1$, then the best response is for firm 1 to charge $m - 1$. If $m > n$, then firm 1's regret is $(m - 1 - n)100$; if $m = n > 1$, firm 1's regret is $(n/2 - 1)100$; if $m = n = 1$, firm 1's regret is 0; and if $m < n$, firm 1's regret is $(m - 1)100$. If $m = 1$, firm 1's best response is to charge 1, and the regret is 0 if $n = 1$ and 100 if $n > 1$. It follows that firm 1's regret is $\max((199 - n)100, (n - 2)100)$. Clearly if $n = 0$, firm 1's regret is 199×100 (if firm 2 charges 200). Thus, firm 1 minimizes regret by playing 100 or 101, and similarly for firm 2. A second round of regret minimization, with respect to $\{100, 101\}$, leads to 100 as the unique strategy that results from iterated regret minimization. This seems far closer to what is done in practice in many cases. ▮

Example 3.10 *The Nash bargaining game [Nash 1950]:* In this 2-player game, each player must choose an integer between 0 and 100. If player 1 chooses x and player 2 chooses y and $x + y \leq 100$, then player 1 receives x and player 2 receives y; otherwise, both players receive 0. All strategy profiles of the form $(x, 100 - x)$ are Nash equilibria. The problem is deciding which of these equilibria should be played. Nash [1950] suggested a number of axioms, for which it followed that $(50, 50)$ was the unique acceptable strategy profile.

Using iterated regret minimization leads to the same result, without requiring additional axioms. Using arguments much like those used in the case of Bertrand competition, it is easy to see that the regret of playing x is $\max(100-x, x-1)$. If the other player plays $y \leq 100-x$, then the best response is $100-y$, and the regret is $100-y-x$. Clearly, the greatest regret is $100-x$, when $y=0$. On the other hand, if the other player plays $y > 100-x$, then the best response is $100-y$, so the regret is $100-y$. The greatest regret comes if $y = 100-x+1$, in which case the regret is $x-1$. It follows that regret is minimized by playing either 50 or 51. Iterating regret minimization with respect to $\{50, 51\}$ leaves us with 50. Thus, $(50, 50)$ is the only strategy profile that survives iterated regret minimization.

We have implicitly assumed here that the utility of a payoff of x is just x. If, more generally, it is $u(x)$, and u is an increasing function, then the same argument shows that the regret is $\max(u(100)-u(x), u(x-1)-u(0))$. Again, there will be either one value for which the regret is maximized (as there would have been above if we had taken the total to be 99 instead of 100) or two consecutive values. A second round of regret minimization will lead to a single value; that is, again there will be a single strategy profile of the form (x, x) that survives iterated regret minimization. However, x may be such that $2x < 100$ or $2x > 100$. This can be viewed as a consequence of the fact that, in a strategy profile that survives iterated regret minimization, a player is not making a best response to what the other players are doing. ∎

We next show that that iterated regret and Nash equilibrium agree on Prisoner's Dilemma. This follows from a more general observation, that iterated regret always recommends a dominant action. A *dominant action* a for player i is one such that $u_i(a, \vec{b}_{-i}) \geq u_i(a', \vec{b}_{-i})$ for all $a' \in A_i$ and $\vec{b} \in A$. We can similarly define a dominant (mixed) strategy. It is easy to see that dominant actions survive iterated deletion of weakly and of strongly dominated actions with respect to A, and are rationalizable. Indeed, if there is a dominant action, the only actions that survive one round of iterated deletion of weakly dominated strategies are dominant actions. Similar observations hold in the case of mixed strategies. The next result shows that iterated regret minimization acts like iterated deletion of weakly dominated strategies in the presence of dominant actions and strategies.

Proposition 3.11 *Let $G = ([n], A, \vec{u})$ be a strategic game. If player i has a dominant action a_i, then*

(a) $\mathcal{RM}_i(A) = \mathcal{RM}_i^\infty(A)$*;*

(b) $\mathcal{RM}_i(A)$ *consists of the dominant actions in G.*

Proof: Every action that is dominant has regret 0 (which is minimal); furthermore, only dominant actions have regret 0. It follows that $\mathcal{RM}_i(A)$ consists of the dominant actions for i (if such actions exist). Since none of these actions will be removed in later deletions, it follows $\mathcal{RM}_i(A) = \mathcal{RM}_i^\infty(A)$. ∎

Example 3.12 *(Repeated) Prisoner's Dilemma:* Recall that in Prisoner's Dilemma, players can either cooperate (c) or defect (d). They payoffs are $u(d, d) = (u_1, u_1))$,

$u(c,c) = (u_2, u_2)$, $u(d,c) = (u_3, 0)$, $u(c,d) = (0, u_3)$, where $0 < u_1 < u_2 < u_3$ and $u_2 > u_3/2$ (so that alternating between (c,d) and (d,c) is not as good as always cooperating). It is well known (and easy to check) that d is the only dominant action for both 1 and 2, so it follows by Proposition 3.11 that traditional solutions concepts coincide with iterated regret minimization for this game.

Things get more interesting if we consider repeated Prisoner's Dilemma. We show that if both players use iterated regret deletion, they will defect in every round, both in finitely and infinitely repeated Prisoner's Dilemma.

First consider Prisoner's Dilemma repeated n times. Let s_{ad}, the strategy where player 1 always defects, and let S consist of all pure strategies in n-round Prisoner's Dilemma.

Lemma 3.13 *$regret_1^S(s_{ad}) = (n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$. Moreover, if s is a strategy for player 1 where he plays c before seeing player 2 play c (i.e., where player 1 either starts out playing c or plays c at the kth for $k > 1$ move after seeing player 2 play d for the first $k-1$ moves), then $regret_1^S(s) > (n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$.*

It follows from Lemma 3.13 that the only strategies that remain after one round of deletion are strategies that start out defecting, and continue to defect as long as the other player defects. If the players both play such a strategy, they both defect at every round. Thus, all these strategies that survive one round of deletion survive iterated deletion. It follows that with iterated regret minimization, we observe defection in every round of finitely repeated prisoners dilemma. Essentially the same argument shows that this is true in infinitely repeated Prisoner's Dilemma (where payoffs are discounted by δ, for $0 < \delta < 1$). By way of contrast, while always defecting is the only Nash equilibrium in finitely repeated Prisoner's Dilemma, the Folk Theorem shows that for all p with $0 \le p \le 1$, if δ is sufficiently close to 1, there is an equilibrium in which p is the fraction of times that both players cooperate. Thus, with Nash equilibrium, there is a discontinuity between the behavior in finitely and infinitely repeated Prisoner's Dilemma that does not occur with regret minimization. Nevertheless, intuition suggests that there should be a way to justify cooperation using regret minimization, just as in the case of the Centipede Game. This is indeed the case, as we show in Section 3.5. ∎

Example 3.14 *Hawk-Dove:*[6] In this game, $A_1 = A_2 = \{d, h\}$; a player can choose to be a dove (d) or a hawk (h). The payoffs are something like Prisoner's Dilemma (with h playing the role of "defect"), but the roles of a and 0 are switched. Thus, we have $u(d,d) = (b,b), u(d,h) = (a,c), u(h,d) = (c,a), u(h,h) = (0,0)$, where again $0 < a < b < c$. This switch of the role of a and 0 results in the game having two Nash equilibria: (h,d) and (d,h). But d is the only action that minimizes regret (yielding regret $c - b$). Thus, $\mathcal{RM}(A) = \mathcal{RM}^\infty(A) = (d,d)$. ∎

[6] This game is sometimes known as *Chicken*.

Example 3.15 In all the examples considered thus far, there were no strategies that were strongly dominated by pure strategies, so if we restrict to pure strategies, all strategies survived iterated deletion of strongly dominated strategies. The game described below shows that the strategies that survive iterated deletion of strongly dominated strategies by pure strategies can be disjoint from those that survive iterated regret minimization.

	x	y
a	$(0,100)$	$(0,0)$
b	$(1,0)$	$(1,1)$

First consider iterated deletion of strongly dominated strategies. For player 1, b strongly dominates a. Once a is deleted, y strongly dominates x for player 2. Thus, iterated deletion of strongly dominated strategies leads to the unique strategy profile (b, y). Now consider regret minimization. The regret of x is less than that of y, while the regret of b is less than that of a. Thus, iterated regret minimization leads to (b, x). ∎

Note that in the examples above, the deletion process converges after two steps. We can construct examples of games where we need $\max(|A|_1 - 1, \ldots, |A|_n - 1)$ deletion steps. The following example shows this in the case that $n = 2$ and $|A_1| = |A_2|$, and arguably illustrates some problems with iterated regret minimization.

Example 3.16 Consider a symmetric 2-player game with $A_1 = A_2 = \{a_1, \ldots, a_n\}$. If both players play a_k, then the payoff for each one is k. For $k > 1$, if one player plays a_k and the other plays a_{k-1}, then the player who plays a_k get -2, and the other gets 0. In all other cases, both players get payoff of 0. The ij entry of the following matrix describes player h's payoff if h plays a_i and player $2 - h$ plays a_j:

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ -2 & 2 & \ldots & 0 & 0 \\ 0 & -2 & 3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & -2 & n \end{bmatrix}$$

Note that a_n has regret $n + 1$ (if, for example player 1 plays a_n and player 2 plays a_{n-1}, player 1 could have gotten $n + 1$ more by playing a_{n-1}; all other actions have regret n. Thus, only a_n is eliminated in the first round. Similar considerations show that we eliminate a_{n-1} in the next round, then a_{n-2}, and so on. Thus, a_1 is the only pure strategy that survives iterated regret minimization.

Note that (a_k, a_k) is a Nash equilibrium for all k. Thus, the strategy that survives iterated regret minimization is the one that is Pareto dominated by all other Nash equilibria. We get a similar result if we modify the payoffs so that if both players play

a_k, then they both get 2^k, while if one player plays a_k and the other plays a_{k-1}, then the one that plays a_k get $-2-2^{k-1}$, and the other gets 0. Suppose that $n = 20$ in the latter game. Would players really accept a payoff of 2 when they could get a payoff of over 1,000,000 if they could coordinate on a_{20}? Perhaps they would not play a_{20} if they were concerned about the loss they would face if the other player played a_{19}.

The following variant of a *generalized coordination game* demonstrates the same effect even without iteration.

	a	b
a	$(1,1)$	$(0,-10)$
b	$(-10,0)$	$(10,10)$

Clearly (b,b) is the Pareto optimal Nash equilibrium, but playing b has regret 11, whereas a has regret only 10; thus (a,a) is the only profile that minimizes regret. Note, however, that (a,a) is the *risk dominant* Nash equilibrium. (Recall that in a *generalized coordination game*—a 2-player, 2-action game where $u_1(a,a) > u_1(b,a)$, $u_1(b,b) > u_1(a,b)$, $u_2(a,a) > u_2(a,b)$, and $u_2(b,b) > u_2(b,a)$—the Nash equilibrium (a,a) is risk dominant if the product of the "deviation losses" for (b,b) (i.e., $(u_1(b,a) - u_1(b,b))(u_2(a,b) - u_2(b,b))$) is higher than the product of the deviation losses for (a,a).) In the game above, the product of the deviation losses for (b,b) is $100 = (0-10)(0-10)$, while the product of the deviation losses for (a,a) is $121 = (-10-1)(-10-1)$; thus, (a,a) is risk dominant. In fact, in every generalized coordination game, the product of the deviation losses for (x,x) is $\mathit{regret}_1(x)\mathit{regret}_2(x)$, so if the game is symmetric (i.e., if $u_1(x,y) = u_2(y,x)$, which implies that $\mathit{regret}_1(x) = \mathit{regret}_2(x)$), the risk dominant Nash equilibrium is the only action profile that minimizes regret. (It is easy to see that this is no longer the case if the game is asymmetric.) ∎

3.4.2 Mixed strategies

Applying regret in the presence of mixed strategies can lead to quite different results than if we consider only pure strategies. We first generalize Proposition 3.11, to show that if there is a dominant strategy that happens to be pure (as is the case, for example, in Prisoner's Dilemma), nothing changes in the presence of mixed strategies. But in general, things can change significantly.

Proposition 3.17 *Let $G = ([n], A, \vec{u})$ be a strategic game. If player i has a dominant action a_i, then $\mathcal{RM}_i(\Sigma) = \mathcal{RM}_i^\infty(\Sigma) = \Delta(\mathcal{RM}_i(A))$.*

Proof: The argument is similar to the proof of Proposition 3.11 and is left to the reader. ∎

To understand what happens in general, we first shows that we need to consider regret relative to only pure strategies when minimizing regret at the first step. (The proof is relegated to the Appendix.)

Proposition 3.18 *Let $G = ([n], A, \vec{u})$ be a strategic game and let σ_i be a mixed strategy for player i. Then $\mathit{regret}_i^\Sigma(\sigma_i) = \max_{\vec{a}_{-i} \in A_{-i}} \mathit{regret}^{\mathcal{S}_i}(\sigma_i \mid \vec{a}_{-i})$.*

Example 3.19 *Roshambo (Rock-Paper-Scissors):* In the rock-paper-scissors game, $A_1 = A_2 = \{r, s, p\}$; rock (r) beats scissors (s), scissors beats paper (p), and paper beats rock; $u(a, b)$ is $(2, 0)$ if a beats b, $(0, 2)$ if b beats a, and $(1, 1)$ if $a = b$. If we stick with pure strategies, by symmetry, we have $\mathcal{RM}_1(A) = \mathcal{RM}_2(A) = \mathcal{RM}_1^\infty(A) = \mathcal{RM}_2^\infty(A) = \{r, s, p\}$. If we move to mixed strategies, it follows by Proposition 3.18 that picking r, s, and p each with probability 1/3 is the only strategy that minimizes regret. (This is also the only Nash equilibrium.) ∎

Example 3.20 *Matching pennies with mixed strategies:* Consider again the matching pennies game, where $A_1 = A_2 = \{a, b\}$, and $u(a, a) = u(b, b) = (80, 40), u(a, b) = u(b, a) = (40, 80)$. Recall that $\mathcal{RM}_1^\infty(A) = \mathcal{RM}_2^\infty(A) = \{(a, b)\}$. Consider a mixed strategy that puts weight p on a. By Proposition 3.18, the regret of this strategy is $\max(40(1 - p), 40p)$, which is minimized when $p = \frac{1}{2}$ (yielding regret 20). Thus randomizing with equal probability over a, b is the only strategy that minimizes regret; it is also the only Nash equilibrium. But now consider the asymmetric matching pennies game, where $u(a, a) = (320, 40)$. Recall that $\mathcal{RM}^\infty(A) = (a, b)$. Since the utilities have not changed for player 2, it is still the case that $1/2a + 1/2b$ is the only strategy that minimizes regret for player 2. On the other hand, by Proposition 3.18, the regret of the strategy for player 1 that puts weight p on a is $\max(280(1-p), 40p)$, which is minimized when $p = .875$. Thus $\mathcal{RM}^\infty(\Sigma) = (.875a + .225b, 0.5a + 0.5b)$. ∎

Example 3.21 *Coordination games with mixed strategies:* Consider again the coordination game where $A_1 = A_2 = \{a, b\}$ and $u(a, a) = (k, k), u(b, b) = (1, 1), u(a, b) = u(b, a) = 0$. Recall that if $k > 1$, then $\mathcal{RM}_1^\infty(A) = \mathcal{RM}_2^\infty(A) = \{(a)\}$, while if $k = 1$, then $\mathcal{RM}_1^\infty(A) = \mathcal{RM}_2^\infty(A) = \{a, b\}$. Things change if we consider mixed strategies. Consider a mixed strategy that puts weight p on b. By Proposition 3.18, the regret of this strategy is $\max(kp, 1 - p)$ which is minimized when $p = \frac{1}{k+1}$ (yielding regret $\frac{k}{k+1}$). Thus, mixed strategies that minimize regret can put positive weight on actions that have sub-optimal regret. ∎

Example 3.22 *Traveler's Dilemma with mixed strategies:* As we saw earlier, each of the choices 96–100 has regret 3 relative to other pure strategies. It turns out that there are mixed strategies that have regret less than 3. Consider the mixed strategy that puts probability 1/2 on 100, 1/4 on 99, and decreases exponentially, putting probability $1/2^{98}$ on both 3 and 2. Call this mixed strategy σ. Let Σ consist of all the mixed strategies for Traveler's Dilemma.

Lemma 3.23 $\mathit{regret}_1^\Sigma(\sigma) < 3$.

The proof of Lemma 3.23 shows that $\mathit{regret}_1^\Sigma(\sigma)$ is not much less than 3; it is roughly $3 \times (1 - 1/2^{101-k})$. Nor is σ the strategy that minimizes regret. For example, it follows from the proof of Lemma 3.23 that we can do better by using a strategy that puts probability 1/2 on 98 and decreases exponentially from there. While we have not computed the exact strategy that minimizes (or strategies that minimize) regret—the computation is nontrivial and does not add much insight—we can make two useful observations:

- The mixed strategy that minimizes regret places probability at most $3/(99-k)$ on the pure strategies k or less. For suppose player places probability α on the pure strategies k or less. If player 2 plays 100, player 1 could have gotten 101 by playing 99, and gets at most $k+2$ by playing k or less. Thus, the regret is at least $(99-k)\alpha$, which is at least 3 if $\alpha \geq 3/(99-k)$. Thus, for example, the probability that 90 or less is played is at most 1/3.

- The strategy that minimizes regret has regret greater than 2.9. To see this, note that the strategy can put probability at most 3/97 on 2 and at most 3/96 on 3. This means that the regret relative to 3 is at least

$$(1-3/96-3/97)3+3/96 = 3-6/96-3/97 > 2.9).$$

The fact that it is hard to compute the exact strategy the minimizes regret suggests that people are unlikely to be using it. On the other hand, it is easy to compute that the optimal strategy puts high weight on actions in the high 90's. In retrospect, it is also not surprising that one can come close to minimizing regret by putting some weight on (almost) all actions. This was also the case in Example 3.21; as we observed there, we can sometimes do better by putting some weight even on actions that do not minimize regret.[7] Interestingly, the distribution of strategies observed by Becker, Carter, and Naeve [2005] is qualitatively similar to the distribution induced by a mixed strategy that is close to optimal. If everyone in the population was playing a mixed strategy that was close to optimal in terms of minimizing regret, we would expect to see something close to the observed distribution. ∎

The phenomena observed in the previous two examples apply to all the other examples considered in Section 3.4.1. For example, in the Bertrand competition, while the pure strategies of least regret (100 and 101) have regret 9,900, there are mixed strategies with regret less than 7,900 (e.g., by putting probability 1/5 on each of 80, 100, 120, 140, and 160). We can do somewhat better than this, but not much. Moreover, we believe that in both Traveler's Dilemma and in Bertrand Competition there is a unique mixed strategy that minimizes regret, so that one round of deletion will suffice. This is not true in general, as the follow example shows.

[7] It is not necessarily the case that the support of the optimal strategy consists of all actions, or even all undominated actions. For example, consider a coordination game with three actions a, b, and c, where $u(a,a) = u(b,b) = k$, $u(c,c) = 1$, and $u(x,y) = 0$ if $x \neq y$. If $k > 2$, then the strategy that minimizes regret places probability 1/2 on each of a and b, and probability 0 on c.

Example 3.24 Consider a 2-player symmetric game where $A_1 = A_2 = \{a_{mk} : m = 1, \ldots, n, k = 1, 2\}$. Define

$$u_1(a_{ij}, a_{kl}) = \begin{cases} -3^{max(i,k)} & \text{if } i \neq k \\ 0 & \text{if } i = k,\ j = l \\ -3^{i+1} & \text{if } i = k,\ j \neq l \end{cases}$$

Let Σ consists of all mixed strategies in this game. We claim that, for every strategy σ for player 1, $regret_1^{\Sigma}(\sigma) \geq 3^n$, and similarly for player 2. To see this, consider a mixed strategy of the form $\sum_{ij} p_{ij} a_{ij}$. The best response to a_{nj} is a_{nj}, which gives a payoff of 0. Thus, the regret of this strategy relative to a_{n1} is $3^n(\sum_{ij, i \neq n} p_{ij} + 3p_{n2})$. Similarly, the regret relative to a_{n3} is $3^n(\sum_{ij, i \neq n} p_{ij} + 3p_{n1})$. Thus, the sum of the regrets relative to a_{n1} and a_{n2} is $3^n(2 + p_{n1} + p_{n_2})$. It follows that the regret relative to one of a_{n1} and a_{n2} is at least 3^n. It also easily follows that every convex combination of strategies a_{ij} with $i < n$ has regret exactly 3^n (the regret relative to a_{n1} is 3^n, and the regret relative to every other strategy is no worse). Moreover, every strategy that puts positive weight on a_{n1} or a_{n2} has regret greater than 3^n. Thus, at the first step, we eliminate all and only strategies that put positive weight on a_{n1} or a_{n2}. An easy induction shows that at the kth step we eliminate all and only strategies that put positive weight on $a_{(n-k+1)1}$ or $a_{(n-k+1)2}$. After $n-1$ steps of iterated deletion, the only strategies that are convex combinations of a_{11} and a_{12}. One more step of deletion leaves us $1/2a_{11} + 1/2a_{12}$.[8]

In the case of pure strategies, it is immediate that there cannot be more than $|A_1| + \cdots + |A_n|$ rounds of deletion, although we do not have an example that requires more than $\max(|A_1|, \ldots, |A_n|)$ rounds. Example 3.24 shows that $\max(|A_1|, \ldots, |A_n|)$ may be required with mixed strategies, but all that follows from Theorem 3.3 is that the deletion process converges after at most countably many steps. We conjecture that, in fact, the process converges after at most $\max(|A_1|, \ldots, |A_n|)$ steps, both with pure and mixed strategies, but we have not proved this.

3.5 Iterated regret minimization with prior beliefs

We have assumed that we start the deletion process with all pure (resp., mixed) strategy profiles. Moreover, we have assumed that, at all stages in the deletion process (and, in particular, at the beginning), the set of strategies that the agents consider possible is the same for all agents. More generally, we could allow each agent i could start a stage in the deletion process with a set Σ^i of strategy profiles. Intuitively, the strategies in Σ^i_j are the the only strategies that i is considering for j. For $j \neq i$, it is perhaps most natural to think of Σ^i_j as representing i's beliefs about what strategies j will use;

[8] In this example, at every step but the last step, the set of strategies that remain consist of all convex combinations of a subset of pure strategies. But this is not necessarily the case. If we replace 3^k by 2^k in all the utilities above, then we do not eliminate all strategies that put positive weight on a_{n1} or a_{n2}; in particular, we do not eliminate strategies that put the *same* weight on a_{n1} and a_{n2} (i.e., where $p_{n1} = p_{n2}$).

however, it may also make sense to interpret Σ_j^i as a representative set of j's strategies from i's point of view, or as the only ones that i is considering but it is too complicated to consider them all (see below). For $i = j$, Σ_i^i is the set of strategies that i is still considering using; thus, i essentially ignores all strategies other than those in Σ_i^i. When we do regret minimization with respect to a single set $\mathcal{S}$ of strategy profiles (as we do in the definition of iterated regret minimization), we are implicitly assuming that the players have common beliefs.

The changes required to deal with this generalization are straightforward: each agent simply applies the standard regret minimization operator to his set of strategy profiles. More formally, the generalized regret minimization $\mathcal{RM}'$ takes as an argument a tuple $(\Pi_1, \ldots, \Pi_n)$ of strategy profiles and returns such a tuple; we define $\mathcal{RM}(\Pi_1, \ldots, \Pi_n) = (\mathcal{RM}(\Pi_1), \ldots, \mathcal{RM}(\Pi_n))$.[9]

Example 3.25 *Repeated Prisoner's Dilemma with prior beliefs:* The role of prior beliefs is particularly well illustrated in Repeated Prisoner's Dilemma. In the proof of Lemma 3.13, to show that the regret of a strategy like Tit for Tat is greater $(n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$, it is necessary to consider a strategy where player 2 starts out by defecting, and then cooperates as long as player 1 defects. This seems like an extremely unreasonable strategy for player 2 to use! Given that there are 2^{2^n-1} pure strategies for each player in n-round Prisoner's Dilemma, and computing the regret of each one can be rather complicated, it is reasonable for the players to focus on a much more limited set of strategies. Suppose that each player believes that the other player is using a strategy where plays Tit for Tat for some number k of rounds, and then defects from then on, for some k. Call this strategy s_k. (So, in particular, $s_0 = s_{ad}$ and s_n is Tit for Tat.) Let S_i^* consist of all the strategies s_k for player i; let S_{2-i}^+ be any set of strategies for player $2-i$ that includes S_{2-i}^*. It is easy to see that the best response to s_0 is s_0, and the best response to s_k for $k > 1$ is s_{k-1} (i.e., you are best off defecting just before the other player starts to defect). Thus,

$$regret_i^{S_i^+ \times S_{2-i}^*}(s_k \mid s_l) = \begin{cases} (l-k-1)(u_2-u_1) & \text{if } k < l \\ u_3 + u_1 - 2u_2 & \text{if } k = l > 0 \\ u_3 + u_1 - u_2 & \text{if } k > l > 0 \\ u_1 & \text{if } k > l = 0 \\ 0 & \text{if } k = l = 0. \end{cases}$$

It follows that

$$regret_i^{S_i^+ \times S_{2-i}^*}(s_k) = \begin{cases} \max((n-k-1)(u_2-u_1), u_3 + u_1 - u_2) & \text{if } k \geq 2 \\ \max((n-1)(u_2-u_1), u_3 + u_1 - 2u_2, u1) & \text{if } k = 1 \\ n(u_2 - u_1) & \text{if } k = 0 \end{cases}$$

Intuitively, if player 1 plays s_k and player 2 is playing a strategy in S_2^*, then player 1's regret is maximized if player 2 plays either s_n (in which case 1 would have been better

[9] As we hinted in Section 3.3, an epistemic justification of this more general notion of regret minimization would require a more general notion of lexicographic beliefs, where each player has a separate sequence of beliefs.

off by continuing to cooperate longer) or if player 2 plays s_{k-1} (assuming that $k > 1$), in which case 1 would have been better off by defecting earlier. Thus, the strategy that minimizes player 1's regret is either s_{n-1}, s_1, or s_0. (This is true whatever strategies player 1 is considering for himself, as long as it includes these strategies.) If n is sufficiently large, then it will be s_{n-1}. This seems intuitively reasonable. In the long run, a long stretch of cooperation pays off, and minimizes regret. Moreover, it is not hard to show that allowing mixtures over $s_0, \ldots, s_n$ makes no difference; for large n, s_{n-1} is still the unique strategy that minimizes regret.

To summarize, if each player i believes that the other player $2 - i$ is playing a strategy in S^*_{2-i}—a reasonable set of strategies to consider—then we get a strategy that looks much more like what people do in practice. ∎

Thinking in terms of beliefs makes it easy to relate iterated regret to other notions of equilibrium. Suppose that there exists a strategy profile $\vec{\sigma}$ such that player i's beliefs have the form $\Sigma_i \times \{\sigma_{-i}\}$. That is, player i believes that each of the other players are playing their component of $\vec{\sigma}$, and there are no constraints on his choice of strategy. Then it is easy to see that the strategies that minimize player i's regret with respect to these beliefs are just the best responses to σ_{-i}. In particular, if $\vec{\sigma}$ is a Nash equilibrium, then $(\vec{\sigma}, \ldots, \vec{\sigma}) \in \mathcal{RM}'(\Sigma_1 \times \sigma_{-1}, \ldots, \Sigma_n \times \sigma_{-n})$. The key point here is that if the agent's beliefs are represented by a "small" set, then the agent makes a best response in the standard sense by minimizing regret; minimizing regret with respect to a "large" belief set looks more like traditional regret minimization.

4 Iterated Regret Minimization in 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, $u_i : A \times T \rightarrow I\!R$. 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 $\mathit{regret}_i^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i}, \vec{t})$, is $u_i^{\mathcal{S}_i}(\vec{a}_{-i}, \vec{t}) - u_i(a_i, \vec{a}_{-i}, \vec{t})$. Let $\mathit{regret}_i^{\mathcal{S}}(a_i \mid \vec{t}) = \max_{\vec{a}_{-i} \in \mathcal{S}_{-i}(\vec{t}_{-i})} \mathit{regret}^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i}, \vec{t})$ denote the maximum regret of player i given $\vec{t}$. The *expected regret of a_i given t_i and $\mathcal{S}_{-i}$* is $E[\mathit{regret}_i^{\mathcal{S}_i}(a_i \mid t_i)] = \sum_{\vec{t} \in T} \Pr(\vec{t} \mid t_i) \mathit{regret}_i^{\mathcal{S}_i}(a_i \mid \vec{t})$. Let $\mathit{minregret}^{\mathcal{S}_i}(t_i) = \min_{a_i \in \mathcal{S}_i(t_i)} E[\mathit{regret}_i^{\mathcal{S}}(a_i \mid t_i)]$. We delete all those strategies that do not give an action that minimizes expected regret for each type. Thus, let $\mathcal{RM}_i(\mathcal{S}_i) = \{\sigma \in \mathcal{S}_i : \mathit{regret}_i^{\mathcal{S}_i}(\sigma(t_i) \mid t_i) = \mathit{minregret}^{\mathcal{S}_i}(t_i)\}$, and define $\mathcal{RM}(\mathcal{S}) = \mathcal{RM}_1(\mathcal{S}_1) \times \ldots \times \mathcal{RM}_n(\mathcal{S}_n)$. Having defined the deletion operator, we can apply iterated deletion as before.

Example 4.1 *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 bids. (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 4.2 *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 $v/2 - 1$). ∎

5 Mechanism Design using Regret Minimization

In this section, we show how using regret minimization as the solution concept can help to construct efficient mechanisms. We consider mechanisms where an agent truthfully reporting his type is the unique strategy that minimizes regret, and focus on prior-free mechanisms (i.e., mechanisms that do not depend on the type distribution of players); we call these *regret-minimizing truthful mechanisms*. Additionally, we focus on *ex-post individually-rational* (IR) mechanisms.[10] As the example below shows, regret-minimizing truthful mechanisms can do significantly better than dominant-strategy truthful mechanism.

Example 5.1 *Maximizing revenue in combinatorial auctions:* In a combinatorial auction, there is a set of m indivisible items that are concurrently auctioned to n bidders. The bidders can bid on bundles of items (and have a valuation for each such bundle); the auctioneer allocates the items to the bidders. The standard VCG mechanism is known to maximize social welfare (i.e., the allocation by the auctioneer maximizes the sum of the valuations of the bidders of the items they are assigned), but might yield poor revenue for the seller. Designing combinatorial auctions that provide good revenue guarantees for the seller is a recognized open problem. By using regret minimization as the solution concept, we can provide a straightforward solution.

Consider a combinatorial first-price auction: that is, the auctioneer determines the allocation that maximizes its revenue (based on the bidders' bids), and the winning bidders pay what they bid. Using the same argument as for the the case of a single-item first-price auction, it follows that if a bidder's valuation of a bundle is v (where v is an even number), bidding $v/2$ is the unique bid on that bundle that minimizes his regret. Thus, in a combinatorial first-price auction, the seller is guaranteed to receive $MSW/2$, where MSW denotes the maximal social welfare, that is, the maximum possible sum of the bidders' valuation for an allocation. Clearly MSW is the most that the seller can receive, since a rational bidder would not bid more than his valuation. To additionally get a truthful auction with the same guarantee, change the mechanism so that the winning bidder pay $b/2$ if he bids b; it immediately follows that a bidder with valuation v for a bundle should bid v. (The mechanism is also trivially IR as players never pay more than their valuation.) This should be contrasted with the fact that there is no dominant-strategy implementation (i.e., no mechanism where bidding the valuation

[10] Recall that a mechanism is ex-post individually rational if a player's utility of participating is no less than that of not participating, no matter what the outcome is.

maximizes utility no matter what the other player bid) that guarantees even a positive fraction of MSW as revenue. In fact, as we now show, to guarantee a fraction r of MSW the minimum regret needs to be "large". By way of contrast, dominant strategies have regret 0.

Lemma 5.2 *An efficient, IR, regret-minimizing truthful mechanism that guarantees the seller a fraction r of MSW as revenue has a minimum regret of at least $rMSW - 1$. (In particular, a mechanism that guarantees a revenue of $MSW/2$ must have a minimum regret of at $MSW/2+1$, just like the first-price auction.)*

Proof: The claim already holds if there is a single object and two buyers. Assume by way of contradiction that there exists an efficient, IR, truthful auction where the seller's revenue is at least $rMSW$. Since the auction is efficient and truthful, the bidder with the higher bid b will win the auction. It follows that this bidder must pay at least rb (or else either the revenue guarantee could not be satisfied, or the auction is not truthful), but at most b (or else the auction cannot be both truthful and IR). Thus, player 1's regret when bidding its valuation v is at least $rv-1$, since if player 2 bids 0, player 1 needs to pay at least rv, whereas he could have paid at most 1 by bidding 1 (since with a truthful, IR mechanism, a player will never pay more than he bids). ∎

The following result shows that no regret-minimizing truthful mechanism can do significantly better than the first-price auction in terms of maximizing revenue.

Lemma 5.3 *No efficient, IR, regret-minimizing truthful mechanism can guarantee the seller more than $((\sqrt{5}-1)/2)MSW$ of revenue.*

Proof: As in Lemma 5.2, consider a mechanism for a single object case with two buyers that has a revenue guarantee of $rMSW$. We claim that if player 1 has valuation v, then his regret if he bids αv is at most $\max(\alpha v, v - r\alpha v)$. To see this, note that if player 2 bids $b \leq \alpha v$, then player 1 pays at most αv (by IR and truthfulness), which potentially could have been saved. Thus, his regret is at most αv if player 2 bids less than αv. If, on the other hand, player 2 bids $b > \alpha v$, then player 1 needs to pay at least $r\alpha v$ to win the object (by truthfulness and the revenue guarantee), so his regret is at most $v - r\alpha v$. It is easy to see that $\alpha v = v - r\alpha v$ if $\alpha = 1/(r+1)$. Thus, if $\alpha = 1/(r+1)$, player 1's regret is at most $v/(r+1)$. (We are ignoring here the possibility that $v/(r+1)$ is not an integer, hence not a legal bid. As we shall see, this will not be a problem.) But, by Lemma 5.2, player 1's regret when his valuation is v can be as high as $rv-1$. Thus, we must have $v/(r+1) \geq rv-1$, or equivalently, $(r - 1/(r+1))v \leq 1$. This can be guaranteed for all v only if $r - 1/(r+1) < 0$, so we must have $r < (\sqrt{5}-1)/2$. ∎

6 Related Work

While the notion of regret has been well studied in the context of decision theory (see [Hayashi 2008b] 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. In the words of Linhart and Radner, "in the application to sealed-bid bargaining, the instruction 'Bid so as to minimize your maximum regret' has several advantages over the instruction 'Bid the best response to your opponent's bid.' ". In the computer science literature, 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.[11] 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 $\mathit{regret}_i^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i}, \vec{t})$ as we do (taking $\mathcal{S}_i$ to consist of all strategies for player i), but then define $\mathit{regret}^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$ 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.

Aghassi and Bertsimas [2006] also consider pre-Bayesian games, and use a solution concept in the spirit of that of Hyafil and Boutilier. However, rather than using minimax regret, they use *maximin*, where a maximin action is one with the best worst-case payoff, taken over all the types of the other agents. Just as with the Hyafil-Boutilier notion, the Aghassi-Bertsimas notion collapses to Nash equilibrium if there is a single type.

Even closer to our work is a recent paper by Renou and Schlag [2008] (their work was done independently of ours). Just as we do, they focus on strategic games. Their motivation for considering regret, and the way they do it in the case of pure strategies, is identical to ours (although they do not iterate the deletion process). They allow prior beliefs, as in Section 3.5, and require that these beliefs are described by a closed, convex set of strategies. They are particularly interested in strategy profiles $\vec{\sigma}$ that minimize regret for each agent with respect to all the strategy profiles in an ϵ neighborhood of $\vec{\sigma}$. 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. In particular, they have no analogue to our analysis of purestrategies in Section 3.4.1. If we consider regret relative to the set of all mixed strategy profiles, then we are just in the setting of Section 3.4.2. However, their definition of regret for mixed strategies is somewhat different from ours; see the full paper for a detailed comparison. Appendix C where we also show that, despite these differences, the approaches often give the same results (and are certainly inspired by similar intuitions).

[11] 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 to deal with the more general setting is straightforward.

7 Discussion

The need to find solution concepts that reflect more accurately how people actually play games has long been recognized. This is a particularly important issue because the gap between "descriptive" and "normative" is particularly small in game theory. An action is normatively the "right" thing to do only if it is the right thing to do with respect to how others actually play the game; thus, a good descriptive theory is an essential element of a good normative theory.

There are many examples in the literature of games where Nash equilibrium and its refinements do not describe what 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 1999]). There seems to be some empirical truth to this assumption; for example, 10 of 45 game theorists that submitted pure strategies in the experiments of Becker, Carter, and Naeve [2005] submitted 100. With an assumption of altruism, then the strategies of many of the remaining players can be explained as best responses to their (essentially accurate) beliefs. 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).

While we have illustrated some of the properties of iterated regret minimization, we view this paper as more of a "proof of concept". There are clearly many issues we have left open. We mention a few of the issues we are currently exploring here.

- As we observed in Section 3.5, some behavior is well explained by assuming that agents start the regret minimization procedure with a subset of the set of all strategy profiles, which can be thought of as representing the strategy profiles that the agent is considering. But we need better motivation for where this set is coming from.

- We have considered "greedy" deletion, where all strategies that do not minimize regret are deleted at each step. We could instead delete only a subset of such strategies at each step of deletion. It is well known that if we do this with iterated deletion of weakly dominated strategies, the final set is strongly dependent on the order of deletion. The same is true for regret minimization. Getting an understanding of how robust the deletion process is would be of interest.

- We have focused on normal-form games and Bayesian games. It would also be interesting to extend regret minimization to extensive-form games. A host of new issues arise here, particularly because, as is well known, regret minimization is not time consistent (see [Hayashi 2008a] for some discussion of the relevant issues).
- A natural next step would be to apply our solution concepts to mechanism design beyond just auctions.

A An Epistemic Characterization Using Kripke Structures

Let $G = ([n], A, \vec{u})$ be a strategic game, and let $\mathcal{S}$ denote the full set of mixed strategies. A *(lexicographic belief) Kripke structure* for G has the form $(W, \mathbf{s}, \mathcal{B}_1, \ldots, \mathcal{B}_n)$, where $\mathbf{s}$ associates with each world $w \in W$ a pure strategy profile $\mathbf{s}(w)$ in the game Γ, and $\mathcal{B}_i$ associates with each world w a sequence $(W_0, W_1, \ldots)$ of sets of worlds. For convenience, we use $\mathcal{B}_i^k(w)$ to denote the set W_k in the sequence, so $\mathcal{B}_i(w) = (\mathcal{B}_i^0(w), \mathcal{B}_i^1(w), \ldots)$. Intuitively, $\mathcal{B}_i^0(w)$ consists of the worlds that player i considers most likely at w, the worlds in $\mathcal{B}_i^1(w)$ are less likely, and so on. Thus, the sequence $\mathcal{B}_i(w)$ models player i's beliefs at w. We assume that at each world in $\cup_{j=0}^{\infty}\mathcal{B}_i(w)$ player i uses the same strategy and has the same beliefs. That is, if $w' \in \cup_{j=0}^{\infty}\mathcal{B}_i(w)$, then $\mathbf{s}_i(w) = \mathbf{s}_i(w')$, where $\mathbf{s}_i(w)$ denotes player i's strategy in the strategy profile $\mathbf{s}(w)$, and $\mathcal{B}_i(w) = \mathcal{B}_i(w')$. This captures the intuition that a player knows his own strategy and beliefs.

Given a set W' of worlds, let $\mathbf{s}(W')$ denote the set of strategy profiles associated with the worlds in W'; that is, $\mathbf{s}(W') = \{\mathbf{s}(w') : w' \in W'\}$. Note that while the notion of a lexicographic belief sequence defined in Section 3.3 considers a countable sequence $(\mathcal{S}_0, \mathcal{S}_1, \ldots)$ of strategies, a belief sequence here considers a sequence $\mathcal{B}_i(w)$ of sets of worlds. Player i is *rational in world* w if player i's strategy in w is rational with respect to $\langle \mathbf{s}(\mathcal{B}_i^0(w)), \mathbf{s}(\mathcal{B}_i^1(w)), \ldots \rangle$. where rationality with respect to lexicographic sequences is defined as in Section 3.3.

To capture the type of reasoning we want to do in these Kripke structures, we consider a propositional language similar in spirit to that considered in [Halpern and Pass 2009]. We start with the special proposition *true*, a primitive proposition $play_i(a)$ for each player i and action $a \in A_i$ (which intuitively says that player i is playing action a), a primitive proposition RAT_i for each player i (which, intuitively, says that player i is rational), and close off under conjunction, negation, and the family of modal operators O_i^k. Roughly speaking, $O_i^k\phi$ says that "all player i knows at level k is ϕ". We take $play_{-i}(\vec{a}_{-i})$ to be an abbreviation for $\wedge_{j\neq i} play_j(a_j)$, and $play(\vec{a})$ to be an abbrevation for $\wedge_j play_j(a_j)$.

A formula is either true or false at a world in a Kripke structure. As usual, we write $(M, w) \models \phi$ if ϕ is true at world w in Kripke structure M. The semantics for formulas other that $O_i^k\phi$ is given as follows:

- $(M, w) \models$ *true* (so *true* is vacuously true).

- $(M, w) \models play_i(a)$ if $\mathbf{s}_i(w) = a$
- $(M, w) \models RAT_i$ if $\mathbf{s}_i(w)$ is a best response to the strategy sequence $\langle \mathbf{s}(\mathcal{B}_i^0(w)), \mathbf{s}(\mathcal{B}_i^1(w)), \ldots \rangle$.
- $(M, w) \models \neg\phi$ if $(M, \omega) \not\models \phi$.
- $(M, w) \models \phi \wedge \phi'$ iff $(M, \omega) \models \phi$ and $(M, \omega) \models \phi'$.

The "all agent i knows" operator was first considered by Levesque [1990] and later by Halpern and Lakemeyer [2001]. While our definition is similar in spirit to these earlier definitions, it is more syntactic, and follows the lines of the definition used in [Halpern and Pass 2009]. Intuitively, "all agent i knows is ϕ" if agent i knows ϕ and considers possible every formula consistent with ϕ. The phrase "every formula" makes it clear that the meaning of "all i knows" is dependent on what formulas we consider. For simplicity here, we just consider formulas of the form $play_{-i}(\sigma_{-i})$ for $j \neq i$, so "all i knows is ϕ" says that i considers possible all strategy profiles that other players could use, consistent with ϕ.[12] Thus,

- $(M, w) \models O_i^k\phi$ if $(M, w') \models \phi$ for all $w' \in \mathcal{B}_i^k(w)$ and for all strategies $a_{-i} \in A_{-i}$, if there exists some structure $M^\dagger$ and world $w^\dagger$ in $M^\dagger$ such that $(M^\dagger, w^\dagger) \models \phi \wedge play_{-i}(a_{-i})$, then there exists a world $w'' \in \mathcal{B}^k(w)$ such that $(M, w'') \models play_{-i}(a_{-i})$.

Thus, this formula says that if it is consistent with ϕ that the players other than i play a_{-i}, then player i must consider it possible at the kth level that the players play a_{-i}.

Define the formulas ND_i^k inductively, by taking $ND_i^0 = true$ and

$$ND_i^{k+1} = RAT_i \wedge \wedge_{\ell=0}^k O_i^\ell(\wedge_{j\neq i} ND_j^\ell).$$

This formula captures the intuitions that we outlined in the main text. Player i's top-level belief is that the other players could be using an arbitrary strategy; his first-level belief is that the other players are rational, but that they believe that other players could be using an arbitrary strategy; his second-level belief is that the other players are rational and believe that other players are rational and believe that other players are using an arbitrary strategy; and so on.

The following theorem shows that $\wedge_i ND_i^k$ completely characterizes k rounds of iterated regret minimization.

Theorem A.1 *The following are equivalent:*

- *The action profile $\vec{a}$ survives k rounds of iterated regret minimization (i.e., $\vec{a} \in \mathcal{RM}^k(\mathcal{S})$).*

[12] There is no reason to focus just on strategies. In [Halpern and Pass 2009], we also consider a version of the operator where the language includes all possible beliefs of the other players (i.e., all possible types of the other players). We could do the same here, but for ease of exposition, we focus on the simpler language.

- *There exists a structure M and world w in M such that $(M, w) \models play(\vec{a}) \wedge (\wedge_i ND_i^k)$.*

Proof: We proceed by induction on k. The result is trivially true if $k = 0$. Suppose the result is true for k; we prove it for $k+1$. Suppose that $(M, w) \models play(\vec{a}) \wedge (\wedge_i ND_i^k)$. It is almost immediate from the induction hypothesis that $\mathbf{s}_{-i}(\mathcal{B}_i^\ell)) = \mathcal{RM}_{-i}^\ell(\mathcal{S})$ for $\ell = 0, \ldots, k$ and every player i. Since $(M, w) \models play_i(a_i) \wedge RAT_i$, it follows that $a_i \in \mathcal{RM}_i^{k+1}(\mathcal{S})$. Thus, $\vec{a} \in \mathcal{RM}^{k+1}(\mathcal{S})$.

Now suppose that $\vec{a} \in \mathcal{RM}^{k+1}(\mathcal{S})$. We construct a finite structure M^{k+1} as follows. Let W consist of all the worlds of the form $(\vec{a}, \ell, i)$, where $\vec{a} \in \mathcal{S}$, $\ell \in \{0, \ldots, k+1\}$, and i is a player. Define $\mathbf{s}(\vec{a}, \ell, i) = \mathbf{s}(\vec{a}, \ell, i) = \vec{a}$. Finally, define

$$\mathcal{B}_j^h((\vec{a}, i, \ell)) = \begin{cases} \{(\vec{a}', j, \ell-1) : \vec{a}'_{-j} \in \mathcal{RM}_{-j}^h(\mathcal{S}), a_j = a'_j\} & \text{if } 0 \le h \le \ell,\ j \ne i, \ell > 0; \\ \{(\vec{a}', j, 0) : \vec{a}'_{-j} \in \mathcal{S}_{-j}, a_j = a'_j\} & \text{if } h = \ell = 0,\ j \ne i; \\ \{(\vec{a}', i, \ell) : \vec{a}'_{-i} \in \mathcal{RM}_{-i}^h(\mathcal{S}), a_i = a'_i\} & \text{if } 0 \le h \le \ell,\ j = i; \\ \mathcal{B}_j^\ell((\vec{a}, i, \ell)) & \text{if } h > \ell. \end{cases}$$

It is easy to see that if $a_i = a'_i$, then $\mathcal{B}_i^h(\vec{a}, i, \ell) = \mathcal{B}_i^h(\vec{a}', i, \ell)$, and if $j \ne i$ and $a_j = a'_j$, then $\mathcal{B}_j^h(\vec{a}, i, \ell) = \mathcal{B}_j^h(\vec{a}', j, \ell-1)$. It follows that, for all worlds $w, w' \in M^{k+1}$, if $w' \in \cup_{j=0}^\infty \mathcal{B}_i(w)$, then $\mathbf{s}_i(w) = \mathbf{s}_i(w')$ and $\mathcal{B}_i(w) = \mathcal{B}_i(w')$. Thus, M^{k+1} is indeed a lexicographic belief Kripke structure. Now a straightforward induction on h shows that if $\vec{a} \in \mathcal{RM}^h(\mathcal{S})$ and $h \le \ell$, then $(M, (\vec{a}, \ell, j)) \models \wedge_i ND_i^h$. The desired result follows. ∎

B Proofs

We provide proofs of all the results not proved in the main text here. We repeat the statements for the convenience of the reader.

Theorem 3.3: *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 $\mathcal{RM}(\mathcal{RM}^\infty(\mathcal{S})) = \mathcal{RM}^\infty(\mathcal{S})$.*

Proof: We start with the case of pure strategies, since it is so simple. Since $\mathcal{D}(\mathcal{S}) \subseteq \mathcal{S}$ for any deletion operator $\mathcal{D}$ and set $\mathcal{S}$ of strategy profiles, when we have have equality, then clearly $\mathcal{RM}^\infty(A) = \mathcal{RM}^k(A)$. Since A is finite by assumption, after some point we must have equality. Moreover, we have $\mathcal{RM}(\mathcal{RM}^\infty(A)) = \mathcal{RM}^\infty(A)$.

To deal with the general case, we must work a little harder. The fact that $\mathcal{RM}^\infty(\mathcal{S}) = \mathcal{RM}_1^\infty(\mathcal{S}) \times \ldots \times \mathcal{RM}_n^\infty(\mathcal{S})$ is straightforward and left to the reader. To prove the other parts, we first need the following lemma.

Lemma B.1 *Let $\mathcal{S}$ be a nonempty closed set of strategies. Then $\mathcal{RM}(\mathcal{S})$ is closed and nonempty.*

Proof: We start by showing that $regret_i^{\mathcal{S}_i}$ is continuous. First note that $regret_i^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$ is a continuous function of $\vec{a}_i$. By the closedness (and hence compactness) of $\mathcal{S}$ it follows that $regret_i^{\mathcal{S}}(a_i) = \max_{\vec{a}_{-i}\in\mathcal{S}_{-i}} regret^{\mathcal{S}_i}(a_i \mid \vec{a}_{-i})$ is well defined (even though it involves a max). To see that $regret_i^{\mathcal{S}}$ is continuous, suppose not. This means that there exist some a, δ such that for all n, there exists an a_n within $1/n$ of a such that $|regret_i^{\mathcal{S}}(a_n) - regret_i^{\mathcal{S}}(a)| > \delta$. By compactness of $\mathcal{S}$, it follows by the Bolzano-Weierstrass theorem [Rudin 1976] that there exist a convergent subsequence $(a_{n_m}, regret_i(a_{n_m}))$ which converges to (a, b). We have $|b - regret_i(a)| \geq \delta$, which is a contradiction.

Now, to see that $\mathcal{RM}(\mathcal{S})$ is nonempty, it suffices to observe that, because $\mathcal{S}$ is compact, and $regret_i^{\mathcal{S}}$ is continuous, for each player i, there must be some strategy σ_i such that $regret_i^{\mathcal{S}}(\sigma_i) = minregret_i^{\mathcal{S}}$. Thus, $\sigma_i \in \mathcal{RM}(\mathcal{S})$.

To show that $\mathcal{RM}(\mathcal{S})$ is closed, suppose that $\langle\sigma^m\rangle_{m=1,2,3,\ldots}$ is a sequence of mixed strategy profiles in $\mathcal{RM}(\mathcal{S})$ converging to σ (in the sense that the probability placed by σ^m on a pure strategy profile converges to the probability placed by σ on that strategy profile) and, by way of contradiction, that $\sigma \notin \mathcal{RM}(\mathcal{S})$. Thus, for some player i, $\sigma_i \notin \mathcal{RM}_i(\mathcal{S})$. Note that, since $\mathcal{RM}(\mathcal{S}) \subseteq \mathcal{S}$, the sequence $\langle\sigma_i^m\rangle_{m=1,2,3,\ldots}$ is in $\mathcal{S}$; since $\mathcal{S}$ is closed, $\sigma \in \mathcal{S}$. Let $minregret_i^{\mathcal{S}} = b$. Since $\sigma_i^m \in \mathcal{RM}_i(\mathcal{S})$, we must have that $regret_i^{\mathcal{S}}(\sigma_i^m) = b$ for all m. Since $\sigma_i \notin \mathcal{RM}_i(\mathcal{S})$, there must exist some strategy profile $\vec{\tau} \in \mathcal{S}$ such that $U_i(\vec{\tau}) - U_i(\sigma_i, \tau_{-i}) = b' > b$. But by the continuity of utility, $\lim_m U_i(\sigma_i^m, \tau) = b'$. This contradicts the assumption that $regret_i^{\mathcal{S}}(\sigma_i^m) = b$ for all m. Thus, $\mathcal{RM}(\mathcal{S})$ must be closed. ∎

Returning to the proof of Proposition 3.3, note that since $\mathcal{S}$ is closed and nonempty, it follows by Lemma B.1 that $\mathcal{RM}^k(\mathcal{S})$ is a closed nonempty set for all k. Additionally, note that $\mathcal{RM}^k(\mathcal{S})$ can be viewed as a subset of the compact set $[0,1]^{|A|}$ (since a probability distribution on a finite set X can be identified with a tuple of numbers in $[0,1]^{|X|}$); it follows that $\mathcal{RM}^k(\mathcal{S})$ is also bounded, and thus compact. Finally, note that the set $\{\mathcal{RM}^k(\mathcal{S}) : k = 1,2,3,\ldots\}$ has the *finite intersection property*: the intersection of any finite collection of its elements is nonempty (since it is equal to the smallest element). The compactness of $\mathcal{S}$ now guarantees that the intersection of all the sets in a collection of closed subsets of $\mathcal{S}$ with the finite intersection property is nonempty [Munkres 2000]. In particular, it follows that $\mathcal{RM}^\infty(\mathcal{S})$ is nonempty.

To see that $\mathcal{RM}^\infty(\mathcal{S})$ is a fixed point of the deletion process, suppose, by way of contradiction, that $\sigma_i \in \mathcal{RM}_i^\infty(\mathcal{S}) - \mathcal{RM}_i(\mathcal{RM}^\infty(\mathcal{S}))$. Let $minregret_i^{\mathcal{RM}^\infty(\mathcal{S})} = b$ and choose $\sigma_i' \in \mathcal{RM}_i^\infty(\mathcal{S})$ such that $regret_i^{\mathcal{RM}^\infty(\mathcal{S})}(\sigma_i') = b$. Since $\sigma_i \notin \mathcal{RM}(\mathcal{RM}^\infty(\mathcal{S}))$, it must be the case that $regret_i^{\mathcal{RM}^\infty(\mathcal{S})}(\sigma_i') = b' > b$. By assumption, $\sigma_i \in \mathcal{RM}^\infty(\mathcal{S})$, so $\sigma_i \in \mathcal{RM}^k(\mathcal{S})$ for all k; moreover, $regret_i^{\mathcal{RM}^k(\mathcal{S})}(\sigma_i) \geq b'$. Since $\sigma_i \in \mathcal{RM}^{k+1}(\mathcal{S})$, it follows that $minregret_i^{\mathcal{RM}^k(\mathcal{S})} \geq b'$. This means that there exists a strategy profile $\vec{\tau}^k \in \mathcal{RM}^k$ such that $U_i(\vec{\tau}^k) - U_i(\sigma_i', \vec{\tau}_{-i}^k) \geq b'$. By the Bolzano-Weierstrass theorem, the sequence of strategies $\langle\vec{\tau}^k\rangle_{k=1,2,\ldots}$ has a convergent subsequence $\langle\vec{\tau}^{k_j}\rangle_{j=1,2,\ldots}$ that converges to some strategy profile $\vec{\tau}$. Since $\tau^k \in \mathcal{RM}^m(\mathcal{S})$ for all $k \geq m$, it must be the case that, except for possibly a finite initial segment, this convergent subsequence is in $\mathcal{RM}^m(\mathcal{S})$. Since $\mathcal{RM}^m(\mathcal{S})$ is closed, $\vec{\tau}$, the limit of the convergent subsequence, is in $\mathcal{RM}^m(\mathcal{S})$ for all $m \geq 1$. Thus, $\vec{\tau} \in \mathcal{RM}^\infty(\mathcal{S})$. Now a simple continuity argument shows that $U_i(\vec{\tau}) - U_i(\sigma'_i, \vec{\tau}_{-i}) \geq b' > b$, a contradiction. ∎

Lemma 3.13: *$regret_1^S(s_{ad}) = (n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$. Moreover, if s is a strategy for player 1 where he plays c before seeing player 2 play c (i.e., where player 1 either starts out playing c or plays c at the kth for $k > 1$ move after seeing player 2 play d for the first $k-1$ moves), then $regret_1^S(s) > (n-1)(u_3-u_2)+\max(-u_1, u_2-u_3)$.*

Proof: Let s_c be the strategy where player 2 starts out playing d and then plays c to the end of the game if player 1 plays c, and plays d to the end of the game if player 1 plays d. We have $regret_1^{S_1}(s_{ad} \mid s_c) = (n-1)(u_3 - u_1) - u_1$: player 1 gets gets nu_1 with (s_{ad}, s_c), and could have gotten $(n-1)u_3$ if he had cooperated on the first move and then always defected.

Let s'_c be the strategy where player 2 starts out playing c and then plays c to the end of the game if player 1 plays c, and plays d to the end of the game if player 1 plays d. It is easy to see that $regret_1^{S_1}(s_{ad} \mid s'_c) = (n-1)(u_3 - u_1) + (u_2 - u_3)$. Thus, $regret_1^S(s_{ad}) \geq (n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$. We now show that $regret_1^S(s_{ad}) = (n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$. For suppose that the regret is maximized if player 2 plays some strategy s, and player 1's best response to s is s'. Consider the first place where the play of (s', s) differs from that of (s_{ad}, s). This must happen *after* a move where player 1 plays c with s'. For as long as player 1 plays d with s', player 2 cannot distinguish s' from s_{ad}, and so does the same thing in response. So suppose that player 1 plays c at move k with s'. If player 2 plays c at step k, player 1 gets a payoff of u_3 with (s_{ad}, s) at step k and a payoff of u_2 with (s', s). Thus, player 1's total payoff with (s_{ad}, s) is at least $(n-1)u_1 + u_3$, while his payoff with (s_{ad}, s) is at most $(n-1)u_3 + u_2$; thus, his regret is at most $(n-1)(u_3 - u_1) + (u_2 - u_3)$. On the other hand, if player 2 plays d with s at step k, then player 1's payoff at step k with (s', s) is 0, while his payoff at step k with (s_{ad}, s) is u_1. Thus, his regret is at most $(n-1)(u_3 - u_1) - u_1$. (In both cases, the regret can be that high only if $k = 1$.)

We next show that if s is a strategy for player 1 where he plays c before seeing player 2 play c, then $regret_1^S(s) > (n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$. Suppose that k is the first move where player 1 plays c despite not having seen c before. If $k = 1$ (so that player 1 cooperates on the first move), let s_d be the strategy where player 2 plays d for the first move, then plays c to the end of the game if player 1 has played d for the first move, and otherwise plays d to the end of the game. It is easy to see that Then $regret_1^{S_1}(s \mid s_d) = (n-1)(u_3 - u_1) + u_1$. On the other hand, if $k > 1$, then the regret $regret_1^{S_1}(s \mid s_c) \geq (n-1)(u_3 - u_1)$. Thus, $regret_1^S(s) > (n-1)(u_3 - u_2) + \max(-u_1, u_2 - u_3)$. ∎

Proposition 3.18: *Let $G = ([n], A, \vec{u})$ be a strategic game and let σ_i be a mixed strategy for player i. Then $regret_i^{\Sigma}(\sigma_i) = \max_{\vec{a}_{-i} \in A_{-i}} regret_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i})$.*

Proof: Note that, for all strategies profiles $\vec{\sigma}_{-i}$, there exists some strategy σ_i^* such that

$U_i^{\Sigma_i}(\vec{\sigma}_{-i}) = \max_{\sigma_i' \in \Sigma_i} U_i(\sigma_i', \vec{\sigma}_{-i}) = U_i(\sigma_i^*, \vec{\sigma}_{-i})$. It follows that

$$\begin{array}{lll} U_i^{\Sigma_i}(\vec{\sigma}_{-i}) & = & U_i(\sigma_i^*, \vec{\sigma}_{-i}) \\ & = & \sum_{\vec{a}_{-i} \in A_{-i}} \vec{\sigma}_{-i}(\vec{a}_{-i}) U_i(\sigma_i^*, \vec{a}_{-i}) \\ & \leq & \sum_{\vec{a}_{-i} \in A_{-i}} \vec{\sigma}_{-i}(\vec{a}_{-i}) U_i^{\Sigma_i}(\vec{a}_{-i}). \end{array}$$

Thus

$$\begin{array}{ll} & \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{\sigma}_{-i}) \\ = & U_i^{\Sigma_i}(\vec{\sigma}_{-i}) - U_i(\sigma_i, \vec{\sigma}_{-i}) \\ = & U_i^{\Sigma_i}(\vec{\sigma}_{-i}) - \sum_{\vec{a}_{-i} \in A_{-i}} \vec{\sigma}_{-i}(\vec{a}_{-i}) U_i(\sigma_i, \vec{a}_{-i}) \\ \leq & \sum_{\vec{a}_{-i} \in A_{-i}} \vec{\sigma}_{-i}(\vec{a}_{-i}) U_i^{\Sigma_i}(\vec{a}_{-i}) - \sum_{\vec{a}_{-i} \in A_{-i}} \vec{\sigma}_{-i}(\vec{a}_{-i}) U_i(\sigma_i, \vec{a}_{-i}) \\ = & \sum_{\vec{a}_{-i} \in A_{-i}} \vec{\sigma}_{-i}(\vec{a}_{-i}) \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i}) \\ \leq & \max_{\vec{a}_{-i} \in A_{-i}} \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i}). \end{array}$$

It follows that

$$\mathit{regret}_i^{\Sigma}(\sigma_i) = \max_{\vec{\sigma}_{-i} \in \Sigma_{-i}} \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{\sigma}_{-i}) = \max_{\vec{a}_{-i} \in A_{-i}} \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i}). \blacksquare$$

Lemma 3.23: *$\mathit{regret}_1^{\Sigma}(\sigma) < 3$.*

Proof: By Proposition 3.18, to compute $\mathit{regret}_1^{\Sigma}(\sigma)$, it suffices to compute $\mathit{regret}_1^{\Sigma_1}(\sigma \mid a)$ for each action a of player 2. If Player 1 plays σ and player 2 player plays 100, then the best response for player 1 is 99, giving him a payoff of 99. The payoff with σ is

$$\begin{array}{ll} & 100 \times 1/2 + 101 \times 1/4 + 100 \times 1/8 + \cdots + 5 \times 2^{-98} + 4 \times 2^{-98} \\ & 102 \times 1/2 + 101 \times 1/4 + 100 \times 1/8 + \cdots + 5 \times 2^{-98} + 4 \times 2^{-98} - 1 \\ = & 4 \times (1/2 + 1/4 + \cdots + 2^{-98} + 2^{-98}) + \\ & 1 \times (1/2 + 1/4 + \cdots + 2^{-98}) + \\ & 1 \times (1/2 + 1/4 + \cdots + 2^{-97}) + \cdots 1/2 - 1 \\ = & 4 + (1 - 2^{-98} + (1 - 2^{-97}) + \cdots + (1 - 1/2) - 1 \\ = & 102 - (1/2 + 1/4 + \cdots + 2^{-98}) - 1 \\ = & 100 + 1/2^{98}, \end{array}$$

so the regret is less than 1. Similarly, if player 2 plays k with $2 \leq k \leq 99$, the best response for player 1 is $k-1$, which would give player 1 a payoff of $k+1$, while the payoff from σ is

$$\begin{array}{l} (k-2)(1/2 + \cdots + 1/2^{100-k}) + k \times 1/2^{101-k} + (k+1) \times 1/2^{102-k} + k \times 1/2^{103-k} + \\ (k-1) \times 1/2^{104-k} + \cdots + 5 \times 1/2^{98} + 4 \times 1/2^{98}. \end{array}$$

Thus, player 1's regret if player 2 plays k is

$$\begin{array}{ll} & 3 \times (1/2 + \cdots + 1/2^{100-k} + 1/2^{105-k}) + 2 \times 1/2^{104-k} + 1 \times (1/2^{101-k} + 1/2^{103-k}) \\ + & 1/2^{106-k}(4 + 5 \times 1/2 + 6 \times 1/4 + \cdots + (k-4) \times 1/2^{98}) \\ = & 3 \times (1/2 + \cdots + 1/2^{100-k} + 1/2^{105-k}) 2 \times 1/2^{104-k} + 1 \times (1/2^{101-k} + 1/2^{103-k}) + 6 \times 1/2^{106-k}) \\ = & 3 \times (1/2 + \cdots + 1/2^{100-k} + 1/2^{104-k}) 2 \times 1/2^{104-k} + 1 \times (1/2^{101-k} + 1/2^{103-k}) \\ \sim & 3 \times (1 - 1/2^{101-k}) < 3. \end{array}$$

$\blacksquare$

C A Comparison to the Renou-Schlag Approach

As we said, the Renou-Schlag definition of regret in the case of mixed strategies is somewhat different from ours. Our definition of $\mathit{regret}_i^{\mathcal{S}_i}(\sigma_i \mid \vec{\sigma}_{-i})$ does not depend on whether $\vec{\sigma}$ consist of pure strategies or mixed strategies (except that expected utility must be used in the case of mixed strategies, rather than utility). By way of contrast, Renou and Schlag [2008] define

$$\mathit{regret}'_i(\sigma_i \mid \vec{\sigma}_{-i}) = \sum_{a_i \in A_i, a_{-i} \in A_{-i}} \sigma_i(a)\vec{\sigma}_{-i}(\vec{a}_{-i})\mathit{regret}_i^{A_i}(a_i \mid \vec{a}_{-i}),$$

where, as before, $\mathit{regret}_i^{A_i}(a_i \mid \vec{a}_{-i})$ denotes the regret of player i relative to the actions A_i. That is, $\mathit{regret}'_i(\sigma_i \mid \vec{\sigma}_{-i})$ is calculated much like the expected utility of $\vec{\sigma}$ to agent i, in terms of the appropriate convex combination regrets for pure strategies. Note that regret'_i is independent of any set $\mathcal{S}_i$ of strategies.

While we view players as actually choosing mixed strategies, intuitively, Renou and Schlag view players as really choosing only pure strategies. A mixed strategy is taken to represent the beliefs of the players regarding other players' strategies. With this viewpoint, to calculate the regret of an action a relative to a mixed strategy $\vec{\sigma}_{-i}$ for the other players, the player should calculate his regret relative to every pure strategy for the other players, and multiply that by the likelihood of that strategy occurring, according to his beliefs (i.e., according to $\vec{\sigma}_{-i}$). Thus, with this viewpoint, we should have

$$\mathit{regret}'_i(\sigma_i \mid \vec{\sigma}_{-i}) = \sum_{a_{-i} \in A_{-i}} \vec{\sigma}_{-i}(\vec{a}_{-i})\mathit{regret}'_i(\sigma_i \mid \vec{a}_{-i}).$$

The fact that $\mathit{regret}'(\sigma_i \mid \vec{a}_{-i})$ should be $\sum_{a_i \in A} \sigma_i(a_i)\mathit{regret}_i^{A_i}(a_i \mid \vec{a}_{-i})$ is noncontroversial. As the following lemma shows, both approaches agree on the regret relative to a pure strategy, although they disagree in general.

Lemma C.1 *If $\vec{\sigma}$ is a mixed strategy and $\vec{a}$ is a pure strategy, then $\mathit{regret}'_i(\sigma_i \mid \vec{a}_{-i}) = \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i})$, but in general, $\mathit{regret}'_i(\sigma_i \mid \vec{\sigma}_{-i}) \neq \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{\sigma}_{-i})$.*

Proof: For the first claim, it suffices to show that $\mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i}) = \sum_{a_i \in A} \sigma_i(a_i)\mathit{regret}^{A_i}(a_i \mid \vec{a}_{-i})$. It is easy to see that there is a pure strategy for player i that is a best response to $\vec{a}_{-i}$; call it a^*. Thus,

$$\begin{array}{ll} & \mathit{regret}_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i}) \\ = & u_i(a^*, \vec{a}_{-i}) - u_i(\sigma_i, \vec{a}_{-i}) \\ = & u_i(a^*, \vec{a}_{-i}) - \sum_{a_i \in A_i} \sigma_i(a_i)u_i(a_i, \vec{a}_{-i}) \\ = & \sum_{a_i \in A_i} \sigma_i(a_i)(u_i(a^*, \vec{a}_{-i}) - u_i(a_i, \vec{a}_{-i})) \quad [\text{since } u_i(a^*, \vec{a}_{-i}) = \sum_{a_i \in A_i} \sigma_i(a_i)u_i(a^*, \vec{a}_{-i})] \\ = & \sum_{a_i \in A_i} \sigma_i(a_i)\mathit{regret}_i^{A_i}(a_i \mid \vec{a}_{-i}). \end{array}$$

For the second part, consider the following game, where $A_1 = A_2 = \{a, b\}$, and player 1's payoffs are given by the following table:

	a	b
a	3	0
b	0	2

Clearly $regret_1^{\Sigma_1}(a \mid a) = regret_1^{\Sigma_1}(b \mid b) = 0$, $regret_1^{\Sigma_1}(a \mid b) = 2$, and $regret_1^{\Sigma_1}(b \mid a) = 3$. Let σ be the strategy $(1/2)a + (1/2)b$. Then an easy computation shows that $regret'(a \mid \sigma) = 1$, while $regret_1^{\Sigma_1}(a \mid \sigma) = 0$ (since a is the best response to σ). ∎

In light of Lemma C.1, it is perhaps not surprising that the two approaches rank strategies differently. What is perhaps surprising is that, as we now show, if we are considering the strategy that minimizes regret with respect to all strategies, then it does not matter which approach we take.

Proposition C.2 *Let $G = ([n], A, \vec{u})$ be a strategic game and let σ_i be a mixed strategy for player i. Then $regret_i^{\Sigma}(\sigma_i) = \max_{\sigma_{-i} \in \Sigma_{-i}} regret_i'(\sigma_i \mid \sigma_{-i})$.*

Proof: By Proposition 3.18, $regret_i^{\Sigma_i}(\sigma_i) = \max_{\vec{a}_{-i} \in A_{-i}} regret_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i})$. It is also easy follows from the definition of $regret_i'$ that $\max_{\sigma_{-i} \in \Sigma_{-i}} regret_i'(\sigma_i \mid \sigma_{-i}) = \max_{\vec{a}_{-i} \in A_{-i}} regret_i'(\sigma_i \mid \vec{a}_{-i})$. Thus, it suffices to show that $regret_i^{\Sigma_i}(\sigma_i \mid \vec{a}_{-i}) = regret_i'(\sigma_i \mid \vec{a}_{-i})$ for all $\vec{a}_{-i} \in A_{-i}$; this follows from Lemma C.1. ∎

Proposition C.2 shows that, for many of the examples in Section 3.4.2, it would not matter whether we had used $regret'$ instead of $regret$. On the other hand, the difference between the two approaches becomes more significant if we consider prior beliefs. Just as we do, Renou and Schlag consider regret minimization not just with respect to all strategies, but with respect to a set of strategies (representing a player's beliefs). When a player is considering his best response, he always considers all possible strategies for himself. Thus, in our language, Renou and Schlag consider generalized regret minimization with respect to profiles of sets of strategy profiles of the form $(\Sigma_1 \times \Pi_{-1}, \ldots, \Sigma_n \times \Pi_{-n})$; that is, each agent i puts no restriction on his own strategies. (They also require that Π_{-i} be a closed, convex set of strategies.) In general, $regret$ and $regret'$ differ with prior beliefs, as the following example shows.

Example C.3 Consider the game from Lemma C.1. Let σ_α denote the strategy $\alpha a + (1-\alpha)b$; let Π_2 be the (closed, convex) set of strategies for player 2 of the form σ_α, where $\alpha \in [1/5, 3/5]$. It is easy to check that

$$regret_1^{\Sigma_1}(\sigma_\alpha \mid \sigma_{\alpha'}) = \begin{cases} 3\alpha' - (3\alpha\alpha' + 2(1-\alpha)(1-\alpha')) & \text{if } \alpha' \geq 2/5 \\ 2(1-\alpha') - (3\alpha\alpha' + 2(1-\alpha)(1-\alpha')) & \text{if } \alpha' < 2/5, \end{cases}$$

since a is the best response to $\sigma_{\alpha'}$ if $\alpha' \geq 2/5$, and otherwise b is the best response. Simplifying this expression, we get

$$regret_1^{\Sigma_1}(\sigma_\alpha \mid \sigma_{\alpha'}) = \begin{cases} 5(1-\alpha)(\alpha' - 2/5) & \text{if } \alpha' \geq 2/5 \\ 5\alpha(2/5 - \alpha') & \text{if } \alpha' < 2/5. \end{cases}$$

If player 1 believes that player 2 is playing a strategy in Π_2 (i.e., $\alpha' \in [1/5, 3/5]$), then the strategy that minimizes $regret_1^{\Sigma_1}$ is $\sigma_{1/2}$. The strategy $\sigma_{1/2}$ has regret /12; another other strategy will have greater regret relative to either $\sigma_{1/5}$ or $\sigma_{3/5}$. (Note that without this restriction on α', $\sigma_{3/5}$ would have minimized $regret_1$.) By way of contrast,

$$regret_1'(\sigma_\alpha \mid \sigma_{\alpha'}) = 2\alpha(1-\alpha') + 3(1-\alpha)\alpha',$$

and the strategy that minimizes $regret_1'$ is $\sigma_{3/5}$. This guarantees that $regret_1'$ is 6/5, no matter what player 2 does; on the other hand, if player 1 chooses σ_α and $\alpha > 3/5$, then $regret_1'(\sigma_\alpha \mid \sigma_{1/5}) > 6/5$, while if $\alpha < 3/5$, then $regret_1'(\sigma_\alpha \mid \sigma_{3/5}) > 6/5$.

References

Aghassi, M. and D. Bertsimas (2006). Robust game theory. *Mathematical Programming, Series B 107*(1–2), 231–273.

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.

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

Blume, L., A. Brandenburger, and E. Dekel (1991). Lexicographic probabilities and choice under uncertainty. *Econometrica 59*(1), 61–79.

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

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A cognitive hierarchy model of games. *Quareterly Journal of Economics 119*, 861–897.

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.

Dufwenberg, M. and U. Gneezy (2000). Price competition and market concentration: an experimental study. *International Journal of Industrial Organization 18*, 7–22.

Goeree, J. K. and C. Holt (2001). Ten little treasures of game theory, and ten intuitive contradictions. *American Economic Review 91*, 1402–1422.

Goeree, J. K., C. Holt, and T. R. Palfrey (2000). Risk averse behavior in asymmetric matching pennies mames. Discussion paper.

Halpern, J. Y. and G. Lakemeyer (2001). Multi-agent only knowing. *Journal of Logic and Computation 11*(1), 41–70.

Halpern, J. Y. and R. Pass (2009). A logical characterization of iterated admissibility. In *Theoretical Aspects of Rationality and Knowledge: Proc. Twelfth Conference (TARK 2009).*

Hayashi, T. (2008a). Dynamic choice with anticipated regret. Unputlished manuscript.

Hayashi, T. (2008b). Regret aversion and opportunity dependence. *Journal of 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. Twentieth Conference on Uncertainty in Artificial Intelligence (UAI 2004)*, pp. 268–277.

Kreps, D. M. and R. B. Wilson (1982). Sequential equilibria. *Econometrica 50*, 863–894.

Levesque, H. J. (1990). All I know: a study in autoepistemic logic. *Artificial Intelligence 42*(3), 263–309.

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

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

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

Monderer, D. and D. Samet (1989). Approximating common knowledge with common beliefs. *Games and Economic Behavior 1*, 170–190.

Munkres, J. (2000). *Topology* (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall.

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

Nash, J. (1950). The barganing problem. *Econometrica 18*, 155–162.

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*. Cambridge, Mass.: MIT Press.

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

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. *Journal of Economic Theory 25*, 92–100.

Rudin, W. (1976). *Principles of Mathematical Analysis* (Third ed.). McGraw-Hill.

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

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory 4*, 25–55.

Tan, T. and S. Werlang (1988). The Bayesian foundation of solution concepts of games. *Journal of Economic Theory 45*(45), 370–391.