Iterated Regret Minimization in Games with Nonoptimal Nash Equilibria

Sergei Balakin

Last version: Sep 11, 2019

Abstract

Introduced by Halpern and Pass (2012), the concept of iterated regret minimization (IRM) provides solutions that are more reasonable than ones offered by Nash equilibrium (NE) for many games of interest, such as the Traveler's Dilemma, the Centipede Game, and many others. For them, we analyzed new data — the IRM approach prescribes the most profitable strategy better than NE. Also, we apply the IRM to the games with continuous strategy sets, such as Colonel Blotto game and Bertrand-Edgeworth duopoly, and obtain reasonable pricing and allocating policies.

Key words: *regret minimization, Nash equilibrium, game theory*

1 Introduction

In game theory, Nash equilibrium is considered to be the most common solution concept. However, in many games this approach and its improvements (see Osborne and Rubinstein 1994) predict outcomes that are far off from observed data and/or do not maximize players' expected payoffs. For example, they seem inappropriate in the Traveler's Dilemma, the Centipede Game, the Bertrand Competition. Are there any other concepts that would properly describe the best strategy for these games (and many others)? By "best strategy" we mean the strategy that maximizes our utility against a random opponent. To answer this question, we first need to mention k-level thinking (Crawford et al. 2013) and cognitive hierarchy theory (Camerer et al. 2004). Unfortunately, these concepts have one considerable shortage: a necessary demand of an initial belief about opponents' actions. Nevertheless, even if we do not have any ex ante information or beliefs concerning our opponent(s), in some one-shot games we can choose the strategy that maximizes our expected payoff using the iterated regret minimization (IRM) concept.

First mentioning of minimizing regret dates back to Savage (1951). Halpern and Pass (2012) applied this approach to one-shot strategic games and claimed that the IRM solution concept prescribes the most profitable strategy in well-known experiments like the Traveler's Dilemma, the Centipede Game, Nash bargaining, and many others. According to the authors, IRM does not operate with common belief or common knowledge, in contrast to many other solution concepts that involve higher and higher levels of belief regarding other players' rationality. In Nash equilibrium, for example, players are supposed to know the strategy of their opponents. In one-shot games, such knowledge may seem unreasonable. Thus, if a player just wants to play optimally in some sense no matter what other players do, IRM concept provides a tool to capture this intuition.

It is also worth mentioning that Renou and Schlag (2008) elaborated their own minimax regret theory independently of Halpern and Pass. While also focusing on strategic games and having the same motivation for considering regret and identical methods in the case of pure strategies, their definition is slightly different. The approach though often provides the same results because the intuition behind that concept is the same.

While the notion of regret has been significantly developed in decision theory (from Savage 1951 to Hayashi 2008), works on applying regret to game theory have been rare. A bargaining problem was considered by Linhart and Radner (1989), and regret minimization solutions for it appeared to be much more reasonable than Nash equilibrium. In computer science, a hybrid solution concept (Nash equilibrium and regret) was examined in pre-Bayesian games, where each agent's utility depends on both the player's type profile and the action profile, but with no probability assigned to types (Hyafil and Boutilier 2004, Aghassi and Bertsimas 2006).

The motivation and contribution of this paper is the following. On the one hand, we want to check the credibility of IRM concept by analyzing fresh experimental data. We consider a general case of the Traveler's Dilemma and two cases of the Centipede Game with exponential and linear payoffs in the same way as Halpern and Pass did, but examine new data from recent papers. These data support the IRM predictions concerning the most profitable strategies (but not the actual ones) better than many Nash equilibrium improvements. On the other hand, we apply the IRM concept to the games with continuous strategy sets that were not considered before — Bertrand-Edgeworth competitions and Colonel Blotto game — and obtain reasonable price and allocation policies.

Literature and historical reviews open some sections explaining the evolution and importance of the following problems.

2 Iterated Regret Minimization

To define iterated regret minimization (IRM), we are going to simplify Halpern and Pass approach. There is no need to use a deletion operator here, since, in most cases, only two-player games are considered, and the algorithm demands no more than two iterations.

Consider a noncooperative two players' game in its normal form:

$$G = (i \in \{1, 2\}; s_i \in S_i; u_i : S_1 \times S_2 \to \mathbb{R}).$$

For each state s_{-i}, let $u_i^*(s_{-i})$ be the best[1] outcome of player i in state s_{-i}:

$$u_i^*(s_{-i}) = \sup_{s_i \in S_i} u_i(s_i, s_{-i}). \tag{1}$$

Denote *regret* of s_i in state s_{-i} as: $\mathit{regret}_i(s_i, s_{-i}) = u_i^*(s_{-i}) - u_i(s_i, s_{-i})$. Then, the minimax-regret decision rule is

$$s_i^{opt} = \arg \inf_{s_i \in S_i} \sup_{s_{-i} \in S_{-i}} \mathit{regret}_i(s_i, s_{-i}).$$

If s_i^{opt} is not a singleton, we can continue this process from the beginning, considering now that $S_i = s_i^{opt}$. Hence, on the next step, we assume that agents play one of the optimal strategies and want to find one that minimizes regret under this assumption.

According to Halpern and Pass, the idea of the IRM approach is to hedge the player's bets by performing reasonably well no matter what the actual state is. "Reasonably well" is thus dictated by the decision rule: this is the act that minimizes regret. Intuitively, this rule is literally trying to minimize the regret that a player would feel if she discovered what the situation actually was: the "I wish I had chosen s' instead of s" feeling.

Note. We would like to underline the difference between the maximin and the minimax regret approach. The maximin criterion maximizes the worst-case profit. In contrast, the minimax regret aims at making a less conservative decision, by minimizing the opportunity cost from making a suboptimal decision.

[1] Of course, the "best" element may not exist, that is why we use sup (inf) instead of max (min) in the definition. However, for games with finite discrete strategy space this terminology is very convenient, so we roughly reserve it for continuous strategy sets as well.

3 The Traveler's Dilemma

This problem is a great example of how close to the best strategy the IRM prediction can be. It was introduced by Basu (1994). We are going to formulate it in a general case (Halpern and Pass 2012).

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 (in natural numbers) between $\underline{x}$ and $\bar{x}$. 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 the lower amount will get $m + r$, while the other traveler will get $m - r$, where r 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.

The only Nash equilibrium (and the only rationalizable strategy profile) here is $(\underline{x}, \underline{x})$. For small p, this result seems unreasonable. Would anyone interested in maximizing profits ever play $\underline{x}$?

Becker, Carter and Neave (2005) asked 51 members of the Game Theory Society to submit a strategy for the game with $\underline{x} = 2, \bar{x} = 100, r = 2$ (Basu's case). The strategy that worked best (in pairwise matchups against all submitted strategies) was 97 with an average payoff \$85.09. The worst average payoff went to those who played 2; it was \$3.92. This Nash equilibrium strategy was played by 3 people of 51, while 95 or higher was submitted by 33 participants.

Capra et al. (1999) arranged a sequence of experiments and discovered that the result heavily depended on r. For small r people tended to play high values at first and keep doing that when the game was repeated. The higher value of r was given, the lower participants started and the faster they converged to playing $\underline{x}$ with time. Note that the Nash equilibrium is insensitive to the choice of r.

Now consider IRM solution for $r \leqslant (\bar{x} - \underline{x})/2$. It includes four steps.

1. Following (1), obtain the best outcome (to simplify notation, in symmetric games, we will perform calculations for the first player by default):

$$u_1^*(s_2) = s_2 - 1 + r, \qquad s_2 \neq \underline{x}; \qquad\qquad u_1^*(\underline{x}) = \underline{x}.$$

2. Obtain regret:

$$regret_1(s_1, s_2) = u_1^*(s_2) - u_1(s_1, s_2) = \begin{cases} 0, & \text{if } s_1 = s_2 = \underline{x}, \\ r - 1, & \text{if } s_1 = s_2 > \underline{x}, \\ r, & \text{if } s_1 > s_2 = \underline{x}, \\ 2r - 1, & \text{if } s_1 > s_2 > \underline{x}, \\ s_2 - s_1 - 1, & \text{if } s_1 < s_2, \end{cases}$$

3. Minimax calculations: $\min_{s_1 \in S_1} \max_{s_2 \in S_2} regret_1(s_1, s_2) = 2r - 1$.

4. Finally, the optimal solutions satisfy $\bar{x} - s_1 - 1 \leqslant 2r - 1$, which implies $\bar{x} - 2r \leqslant s_1^{opt}$.

Assuming that a strategy from $[\bar{x} - 2r, \bar{x}]$ is used by both players, we can iterate this process. Then the desired strategy is $\bar{x} - 2r + 1$ (regret here is $2r - 2$; all others have regret $2r - 1$). In the Becker, Carter, and Neave case, we have exactly the experimental optimal value: $100 - 2 * 2 + 1 = 97$.

Now let's examine the data collected by Cabrera, Capra and Gomez (2007). They designed an experiment that consisted of three cells of one-shot traveler's dilemma games. Participants were recruited from economics courses at the University of Malaga in Spain. In all treatments, subjects were asked to choose a claim between and including 20 and 120; they were told that the earnings would depend on their decisions and the decisions made by the persons randomly matched with them. The reward/penalty parameter for all sessions was equal to 5. The next table summarizes the experimental design (the absence of a particular claim means that nobody played this strategy):

Claim	20	22	30	35	46	65	69	70	75	80
N	2	1	1	1	1	1	1	1	1	1
AEM	24.8	26.2	33	37	45.7	60.4	63.1	63.5	66.6	69.6
Claim	82	100	102	105	110	111	115	119	120	Total
N	2	3	1	2	1	2	3	1	4	**30**
AEM	70.3	79.3	79.6	80.3	**81.5**	81.3	81.4	81.2	80.5	

The median here is 100; the mode is 120; the mean is 86.21. The mean of AEM (average earnings per match) is 67.5. The standard deviation of AEM is 19.1.

We can see that the most profitable strategy is 110, but 111, 115 and 119 are also very close. Theory above gives us the optimal IRM value $s^{opt} = \bar{x} - 2p + 1 = 120 - 2 * 5 + 1 = 111$. Players who follow the Nash equilibrium get the smallest payoff of 24.8.

4 The Centipede Game

This is another well-known example where empirical observations clash with answers provided by traditional solution concepts. The game was first introduced by Rosenthal (1981). To make it consistent, we formulate this game the same way as Halpern and Pass did.

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 ones. 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.

It can be easily shown by backwards induction that the only Nash equilibrium here is when both players choose to stop immediately: at the first move for player 1 and at the second move for player 2. However, first experiments with linear payoffs (McKelvey and Palfrey 1992, Nagel and Tang 1998) clearly demonstrated that participants tend to cooperate at least for a certain number of rounds (but rarely throughout the whole game) and thus get higher payoffs.

Consider two versions of the Centipede Game and obtain optimal strategies for each of them in the IRM framework.

A. Exponential payoffs. Here, the payoff while ending the game at odd-numbered rounds t is $(2^t+1, 2^{t-1}-1)$, and the payoff while stopping at even-numbered rounds is $(2^{t-1}, 2^t)$ (see Fig.1).

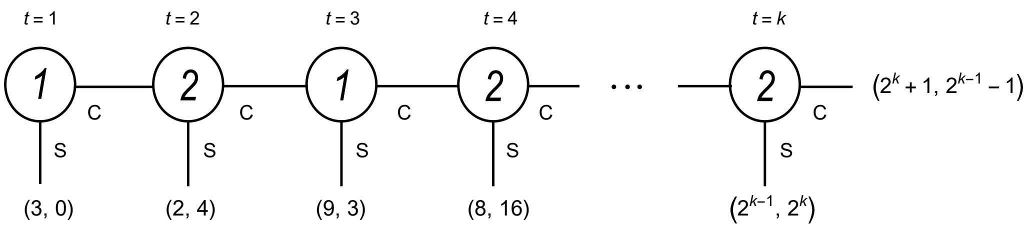

Figure 1: Centipede Game with Exponential Payoffs

Apply the IRM algorithm:

1. From (1) we get[2] (consider k as an even number; for an odd one calculations are similar)

$$u_1^*(t) = 2^{t-1} + 1, \qquad t = 2, 4, \ldots, k, k+2;$$
$$u_2^*(t) = 2^{t-1}, \qquad t = 3, 5, \ldots, k+1, \qquad u_2^*(1) = 0.$$

2. Regret maximization:

$$\max_{s_2 \in S_2} regret_1(s_1, s_2) = \begin{cases} 2^k - 2^{s_1}, \text{ if } s_1 = 1, 3, \ldots, k+1, \\ 1, \text{ if } s_1 = k+1; \end{cases}$$

$$\max_{s_1 \in S_1} regret_2(s_1, s_2) = \begin{cases} 2^{k-1} - 2^{s_2}, \text{ if } s_2 = 2, 4, \ldots, k-2, \\ 1, \text{ if } s_2 = k, k+2, \end{cases}$$

3. Minimizing, we have: $s_1^{opt} = k+1$, $s_2^{opt} = \{k, k+2\}$.

4. Iterating the second time, we finally get the only strategy $(k+1, k)$.

Thus, in the case of exponential payoffs, player 1 continues to play until the very end, and player 2 does the same until the penultimate step.

Note. We used 2 as a base for payoffs, but the argument is correct for any exponential base $a \geqslant 1.325$ (for $k = 4$; for higher k, the value of a can be taken even smaller).

B. Linear payoffs. In this case, the utility of stopping at odd-numbered rounds t is $(t, t-p)$, while the utility of stopping at even-numbered rounds is $(t-p, t)$, where $p > 1$ (see Fig.2).

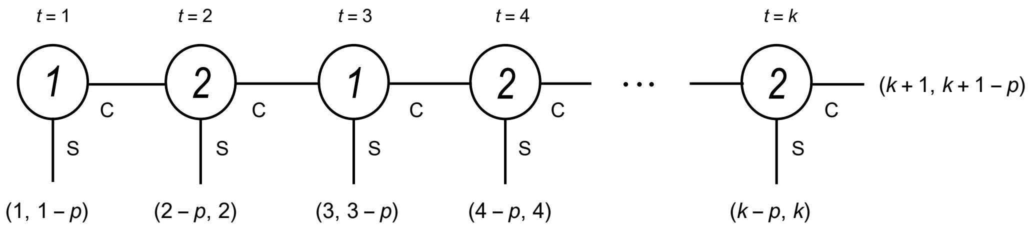

Figure 2: Centipede Game with Linear Payoffs

Since now payoffs increase not so dramatically, we can expect a more relaxed strategy comparing to the exponential case. Agents may not be interested in going on until the very end of the game.

[2]Saying "strategy" in this framework, we mean a *set* of equivalent strategies: all the strategies where player first stops at t (no matter what her actions are afterwards) are payoff equivalent for her.

1. We have $u_i^*(t) = t - 1$ if $i = 1$ and t is even, or if $i = 2$ and t is odd ($t \in \{2, 3, \ldots, k+2\}$; $u_2^*(1) = 1 - p$).

2. Regret maximization:

$$\max_{s_2 \in S_2} regret_1(s_1, s_2) = \begin{cases} k, \text{ if } s_1 = 1, \\ \max\{p - 1; k - s_1 + 1\} \text{ if } s_1 = 3, \ldots, k-1, \\ p - 1, \text{ if } s_1 = k + 1; \end{cases}$$

$$\max_{s_1 \in S_1} regret_2(s_1, s_2) = \begin{cases} k - 2, \text{ if } s_2 = 2, \\ \max\{p - 1; k - s_2\}, \text{ if } s_2 = 4, \ldots, k-2, \\ p - 1, \text{ if } s_2 = k, k+2. \end{cases}$$

3. Finally, we get an optimal regret minimization strategy $(k+1-2m, k-2m)$ for any $1 + 2m \leqslant p < 3 + 2m$ ($m = 0, 1, \ldots, k/2 - 1$). The more severe punishment is, the less patient players are.

Now let's introduce new data from papers of **(BF)** Baghestanian and Frey (2016) and **(ANP)** Atiker, Neilson and Price (2011) and see how it goes along with the model.

(BF): 46 subjects from among the participants at the 28th annual US GO[3] Congress in Black Mountain, NC (August 2012) were recruited. They were offered to play the Centipede Game as described by Nagel and Tang (1998) (see Fig. 3).

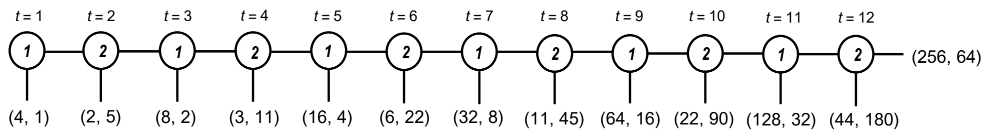

Figure 3: Centipede Game by Nagel and Tang

This is a game with exponential payoffs and $k = 12$. Next table summa-

[3]GO is a very popular East Asian board game, involving very significant strategy with the number of possible games (10^{761} compared, for example, with the estimated 10^{120} possible in chess).

rizes the observations[4]:

Pl. 1 Choice	1	3	5	7	9	11	13	Total
N	0	0	2	5	4	5	8	24
AEM	(4)	(8)	15.4	29.5	53.8	72.4	**97.9**	
Pl. 2 Choice	2	4	6	8	10	12	14	Total
N	0	1	1	2	7	4	7	22
AEM	(5)	11	20.5	33.9	53.4	**71.3**	32.7	

The mean of player 1 and player 2 choice is 10 and 11 correspondingly.

The AEM were calculated for every strategy of player 1(2) matching every strategy of an opponent 2(1). We can see that the most profitable strategy for player 1 is to go on until the end; for player 2 — stop at the last but one node. This is exactly what is predicted by the IRM method. The maximum regrets for player 1 and 2 are:

$$\max_{s_2} \mathit{regret}_1(s_1, s_2) = \{252, 248, \ldots, 128, \mathbf{84}\},$$
$$\max_{s_1} \mathit{regret}_2(s_1, s_2) = \{175, 169, \ldots, 90, \mathbf{74}, 116\},$$

so the minimum value in the first position is approached playing until the last node; in the second position — until the last but one node. The Nash equilibrium strategy happens to be the worst strategy for both players in this framework.

(ANP): The experiment was conducted during the Fall 2009 semester at the University of Tennessee, Knoxville, with a total of 202 undergraduate subjects participated. This Centepede Game is introduced on Fig. 4. This

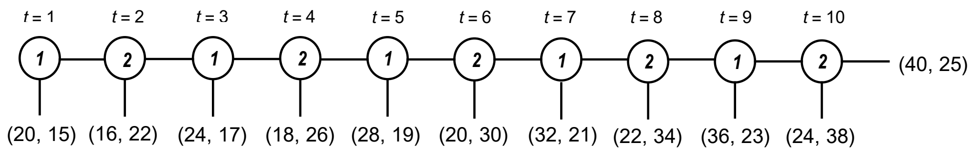

Figure 4: Centipede Game with increasing punishment

is a game with $k = 10$ and linear payoffs, but slight difference from the linear case in section 2 is that in this model punishment also grows linearly.

[4]Values in parentheses mean that they correspond to the average payoff of an "imaginary" player, since no one really plays this strategy.

Observations are summarized in the next table:

Pl. 1 Choice	1	3	5	7	9	11	Total
N	4	27	27	15	17	11	101
AEM	20	23.3	**24.2**	23.7	22.5	21.5	
Pl. 2 Choice	2	4	6	8	10	12	Total
N	9	28	25	20	11	8	101
AEM	21.7	**23.2**	23	22.8	21.3	19.9	

The mean of player 1 and player 2 choice is 5,93 and 6,4 correspondingly.

We can see that the most profitable strategy for player 1 is to stop at the node 5; for player 2 — stop at the node 4. However, the variance here is not as huge as in the exponential case, so one does not lose much stopping at the adjacent node.

Calculate the optimal strategy of IRM method. The maximum regrets for player 1 and 2 are:

$$\max_{s_2} regret_1(s_1, s_2) = \{20, 16, 12, \mathbf{8}, 10, 12\},$$
$$\max_{s_1} regret_2(s_1, s_2) = \{16, 12, \mathbf{8}, 9, 11, 13\}.$$

so the minimum value for the first player is approached playing 7; for the second player — playing 6. It is one node more than introduced by the experiment. The Nash equilibrium strategy happens to be the worst strategy for the first player and far from the best for the second one.

5 The Bertrand duopoly

Although the discrete version of Bertrand competition was already considered by Halpern and Pass, we solve here the original continuous version of the problem just for completeness and to make a bridge to more complicated Bertrand-Edgeworth duopoly.

We consider the classical Bertrand competition where two firms produce a homogeneous product (Bertrand 1883). Here and in the next problem, we assume that each firm has zero production costs.[5] If normalized demand is 1 at any price up to p_M, then payoff for player 1 in this model can be described as follows:

$$u_1(p_1, p_2) = \begin{cases} p_1/2, & \text{if } p_1 = p_2, \\ p_1, & \text{if } p_1 < p_2, \\ 0, & \text{if } p_1 > p_2. \end{cases}$$

[5]This assumption can be made without any loss of generality: it doesn't substantially affect our results.

There is the only Nash equilibrium in this model with zero prices $p_1 = p_2 = 0$ and thus zero profits for both firms. The so-called Bertrand paradox has been criticized as a bad description of the real-life behavior (e.g. nine of ten datasets examined by Slade (1995) considering coffee roasting, banking, wood products etc, and suitable for Bertrand competition rejected the hypothesis of one-shot Nash equilibrium). We'll try to find the optimal strategy of agents in terms of IRM.

The best answer for any strategy $p_2 \in (0, s_M]$ of an opponent is $p_1 = p_2 - \varepsilon$, where ε is an infinitesimal value. Thus, following (1), we have $u_1^*(p_2) = p_2$ (due to continuity of the strategy set), and

$$\sup_{p_2 \in S_2} regret_1(p_1, p_2) = \begin{cases} p_M - p_1, & \text{if } p_1 \leqslant p_M/2, \\ p_1, & \text{if } p_1 > p_M/2. \end{cases}$$

The optimal minimum value can now be easily obtained: $p_1^{opt} = p_M/2$.

Regret minimizing firms in Bertrand duopoly behave in a way that can be explained as a tacit collusion between the agents (explicit cooperation is not allowed). This is economically meaningful; moreover, we did not change neither the model itself no its timing.

6 The Bertrand-Edgeworth duopoly

Here we examine a model of two price setting firms with capacity constraints. It was first introduced by Bertrand (1883) and Edgeworth (1925). Consider the market for a homogeneous good with a demand function $D(p)$ which is assumed to be continuous and strictly decreasing. Each firm has a limited amount of productive capacity S_i, $i = 1, 2$ (not depending on price), such that $D(0) \geqslant S_1 + S_2$. Firms set their own prices p_i, $i = 1, 2$ and cannot cooperate. Thus, the entire market up to its capacity is supplied by the firm that announced the lower price. Another firm serves the residual demand.

Identical consumers choose the lower available price on a first-come-first-serve basis. We assume that the residual demand left for the firm quoting the higher price is a proportion of total demand at that price (Shubik 1955). If firms set the same prices, market gets shared in proportion to their capacities. Thus, the payoff functions of players are:

$$u_i(p_1, p_2) = \begin{cases} p_i \min\{S_i, D(p_i)\}, & \text{if } p_i < p_{-i}, \\ p_i \min\{S_i, \frac{S_i}{S_i + S_{-i}} D(p_i)\}, & \text{if } p_i = p_{-i}, \\ p_i \min\{S_i, \frac{D(p_i)}{D(p_{-i})} \max\{0, D(p_{-i}) - S_{-i}\}\}, & \text{if } p_i > p_{-i}. \end{cases}$$

D'Aspremont and Gabszewicz (1985) showed that the Bertrand-Edgeworth competition does not possess any (pure strategy) Nash equilibrium if $D(p_M) < S_1 + S_2$. They introduced the concept of quasi-monopoly (one capacity is sufficiently bigger than another) that restores the existence of a pseudo equilibrium. As for a mixed-strategy Nash equilibrium, Dasgupta and Maskin (1986) and Dixon (1984) proved just its existence but did not find what it looks like. Also, Allen and Hellwig (1986) showed that the average price set would converge to the competitive one in the case of a large market with many firms.

Let's apply the IRM concept to the Bertrand-Edgeworth duopoly model. Assume for simplicity

$$D(p) = 1 - p, \qquad\qquad S_1 = S_2 = S. \tag{2}$$

Then $S \leqslant 1/2$ and

$$u_1(p_1, p_2) = \begin{cases} p_1 \min\{S, 1-p_1\}, & \text{if } p_1 < p_2, \\ s_1 \min\{S, \frac{1-p_1}{2}\}, & \text{if } p_1 = p_2, \\ p_1 \min\{S, \frac{1-p_1}{1-p_2} \max\{0, 1-p_2-S\}\}, & \text{if } p_1 > p_2. \end{cases} \tag{3}$$

Proposition. In the Bertrand-Edgeworth duopoly with a demand function, capacities and payoffs defined by (2) and (3), with $S \leqslant 1/2$, we have an optimal IRM price value

$$p_1^{opt} = \frac{1}{2}(1 - 3S + \sqrt{1 - 2S + 5S^2}).$$

Proof. Consider three cases: 1) $0 \leqslant S \leqslant 1/4$; 2) $1/4 < S \leqslant 1/3$; 3) $1/3 < S \leqslant 1/2$. Then

$$1) \quad u_1^*(p_2) = \begin{cases} \left(1 - \frac{S(1-p_2)}{1-s_2-S}\right) S, & \text{if } 0 \leqslant p_2 \leqslant 1-2S, \\ p_2 S, & \text{if } 1-2S < p_2 \leqslant 1-S, \\ (1-S)S, & \text{if } 1-S < p_2 \leqslant 1; \end{cases}$$

$$2) \quad u_1^*(p_2) = \begin{cases} \left(1 - \frac{S(1-p_2)}{1-p_2-S}\right) S, & \text{if } 0 \leqslant p_2 \leqslant \frac{1-3S}{1-2S}, \\ \frac{1-p_2-S}{4(1-p_2)}, & \text{if } \frac{1-3S}{1-2S} < p_2 \leqslant \frac{1+4S-\sqrt{1-8S+32S^2}}{8S}, \\ p_2 S, & \text{if } \frac{1+4S-\sqrt{1-8S+32S^2}}{8S} < p_2 \leqslant 1-S, \\ (1-S)S, & \text{if } 1-S < p_2 \leqslant 1; \end{cases}$$

$$3) \quad u_1^*(p_2) = \begin{cases} \frac{1-p_2-S}{4(1-p_2)}, & \text{if } 0 \leqslant p_2 \leqslant \frac{1+4S-\sqrt{1-8S+32S^2}}{8S}, \\ p_2 S, & \text{if } \frac{1+4S-\sqrt{1-8S+32S^2}}{8S} < p_2 \leqslant 1-S, \\ (1-S)S, & \text{if } 1-S < p_2 \leqslant 1. \end{cases}$$

Further, for all three cases, we have

$$f(p_1) = \sup_{p_2}(u_1^*(p_2) - u_1(p_1, p_2)) =$$
$$= \begin{cases} S(1-p_1-S), & \text{if } 0 \leqslant p_1 \leqslant \frac{1}{2}(1-3S+\sqrt{1-2S+5S^2}), \\ p_1(p_1-1+2S), & \text{if } \frac{1}{2}(1-3S+\sqrt{1-2S+5S^2}) < p_1 \leqslant 1-S, \\ (1-S)S, & \text{if } 1-S < p_1 \leqslant 1. \end{cases}$$

Fig. 5 shows the graph of function $f(s_1)$ if $S = 0.3$. The behavior under other values of S is substantially the same.

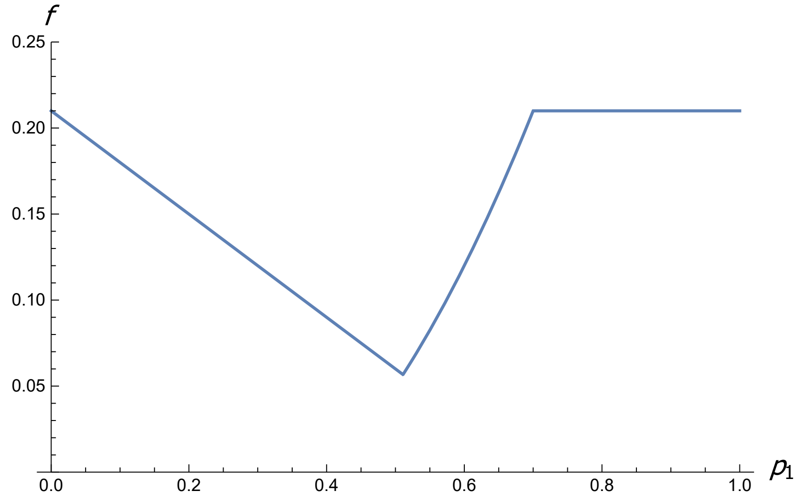

Figure 5: Graph of function $f(p_1)$ if $S = 0.3$

The only minimum value can be reached when $p_1^{opt} = \frac{1}{2}(1-3S+\sqrt{1-2S+5S^2})$.
■

For different values of S, we have a table for corresponding optimal prices:

S	0.025	0.05	0.1	0.2	0.3	0.4	0.5
p^{opt}	0.951	0.903	0.811	0.647	0.511	0.4	0.309
p_M	0.975	0.95	0.9	0.8	0.7	0.6	0.5

These results have a strong economic sense. For any $S \leqslant 1/2$, it is unreasonable for player 1 to make the price less than $1 - 2S$. In this case, there is only a profit loss: if player 2 sets a lower price, she can only serve S of customers, and player 1 serves another S part; if player 2 sets a higher price, player 1 serves S customers immediately. The fact that the optimal value can be a little bit more than $1 - 2S$ follows from the possibility that your opponent sets the price even higher. Of course, the higher price you have, the less this possibility is. For example, in the case $S = 0.1$ we can see a very small deviation 0.011 from the "safe" price 0.8, in the case $S = 0.3$ this deviation 0.111 from the "safe" price 0.4 is quite sufficient. Also, we can see that the ratio p^{opt}/p_M decreases with increasing of S.

Despite the fact that these results are derived for the particular function $D(p)$ and for the symmetric game, the problem can be solved for other types of demand and nonsymmetric firms (for some types of $D(p)$ only numerically). Unfortunately, finding data for a pure one-shot Bertrand-Edgeworth competition is problematic, although there are some for repeated games that demonstrate "Edgeworth cycles" behavior (i.e., see Normann and Fonseca, 2013).

7 Colonel Blotto game

This game was first proposed and solved by Borel in 1921 (and translated into English in 1953) and has been a classic in game theory since then (see Owen, 1968). Here we consider a two-player version where participants need to distribute their limited resources (forces) A and B over two battlefields. We denote players by A and B as well. The commander allocating more forces to a battlefield wins the first battlefield with the payoff α and the second one with the payoff β. The loser does not get anything. If forces are equal, the payoff from that battlefield gets halved. Without loss of generality, we assume that $A > B$ and $\alpha > \beta$. Also consider only the case $A < 2B$ (the solution is trivial otherwise). Let players A and B distribute, respectively, forces a and b on the first battlefield (and thus $A-\alpha$ and $B-\beta$ on the second one). The corresponding total payoffs of the commanders look like

$$u_A = \begin{cases} \beta, & a < b \\ \beta + \alpha/2, & a = b \\ \alpha + \beta, & b < a < b + A - B \\ \alpha + \beta/2, & a = b + A - B \\ \alpha, & a > b + A - B, \end{cases} \qquad u_B = \alpha + \beta - u_A.$$

There exist no pure Nash equilibrium in this game. Indeed, both players face the threat of reducing their payoffs (see Fig. 6). Player A has more resources and thus can always increase her profit by deviation. On the other hand, the maximum payoff $\alpha + \beta$ for player A that may be achieved in the "diagonal" domain $b < a < b + A - B$ is also not ensured: player B can

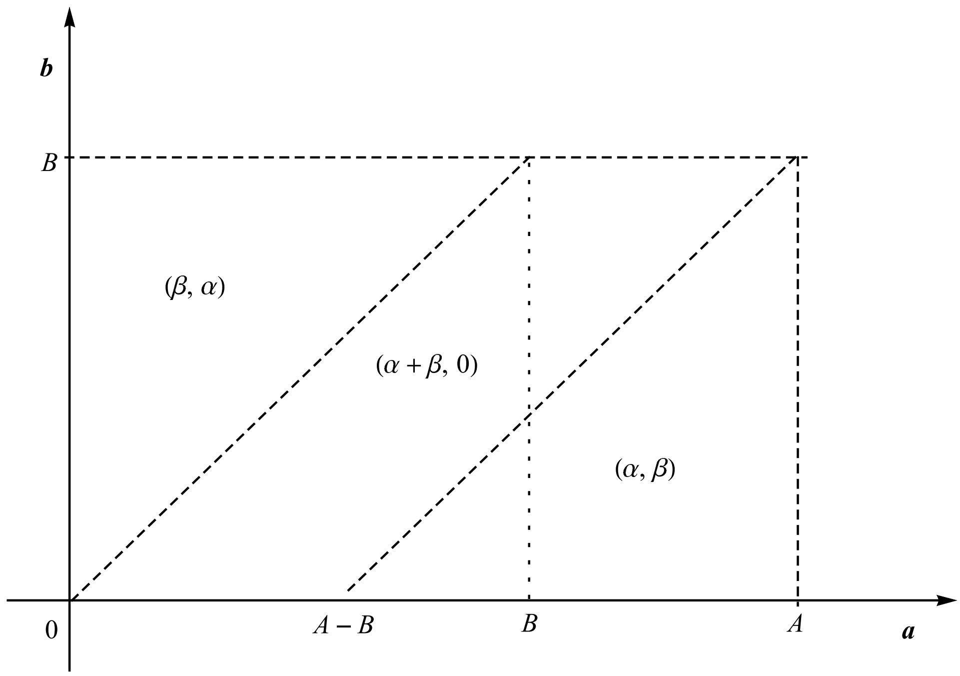

Figure 6: Colonel Blotto Game payoffs

always deviate towards $b = B$ or $b = 0$ to increase her own payoff and reduce the payoff of player A.

However, player A can guarantee her victory on the first battlefield to ensure the payoff α. It can be done just by setting $a > B$ (player B has no available forces to overtake player A on this battlefield). Facing this cautious behavior of a stronger commander, player B can adjust to it and choose the best response $b < 2B - A$ to ensure her victory on the minor battlefield.

Let us consider IRM prediction in this game. For player A, $u_A^*(b) = \sup_a u_A(a, b) = \alpha + \beta$. Then

$$regret_A(a, b) = \begin{cases} \alpha, & a < b \\ 0, & b < a < b + A - B \\ \beta, & a > b + A - B. \end{cases}$$

Thus,

$$f(a) \equiv \sup_b regret_A(a, b) = \begin{cases} \alpha, & a < B \\ \beta, & a > B, \end{cases}$$

and a^{opt} can be anything from the interval $(B, A]$.

Analogous calculations can be done for player B:

$$u_B^* = \alpha, \qquad regret_B(a,b) = \begin{cases} \alpha - \beta, & b < a - (A - B) \\ \alpha, & a - (A-B) < b < a \\ 0, & b > a, \end{cases}$$

$$g(b) \equiv \sup_a regret_B(a,b) = \alpha,$$

and b^{opt} can be anything from $[0, B]$.

Iterating this process again for player B using now restriction $a > B$, we obtain:

$$u_B^* = \beta, \qquad g(b) = \begin{cases} 0, & b < 2B - A \\ \beta, & b < 2B - A \end{cases}$$

and finally $b^{opt} \in [0, 2B - A)$.

Thus, IRM strategy fully corresponds to our prediction. Unfortunately, there is a shortage of literature concerning experiments in Colonel Blotto game. Papers of Modzelewski et al (2009) and Chowdhury et al (2013) consider a sufficiently larger number n of battlefields (five and eight respectively). Although some general patterns are still observed there (i.e., "guerilla warfare" for player B), the behavior becomes different with increasing number of n.

8 Conclusion

There exist many examples of games where Nash equilibrium and its improvements a) do not provide us with the most profitable strategy; b) do not describe well what people really do. Following Halpern and Pass (2012), we have considered a solution concept, iterated regret minimization, that, at least in some games, seems to represent the most profitable strategy better than more standard solution concepts. The majority of these games have undercutting strategies in common, but further work on generalizing the applicable set of games is needed. Also, the IRM notion seems to capture the cooperating and altruistic behavior taking place in many experiments and works relatively well dealing with inexperienced but intelligent players who play a game for the first time. In this one-shot setting, it seems unreasonable to assume that players know strategies of their opponents (as is implicitly stated in Nash equilibrium).

This paper applies IRM method to the two well known problems examined by Halpern and Pass (Traveler's Dilemma and Centipede Game) in a general case and provides some additional experimental data with analysis for them. Also, we examine problems with a continuous number of states (Bertrand-Edgeworth duopoly, Colonel Blotto game) using the IRM concept. In particular, IRM indicates tacit collusion in Bertrand and Bertrand-Edgeworth duopolies with a reasonable price setting. As for Colonel Blotto game, we get well-grounded results that cannot be obtained using standard tools.

A natural next step would be to apply this solution concept to Hotelling games, contests (i.e. Tullock) and auctions in general, and other mechanism design problems. Also, it would be very useful to determine conditions on games under which the IRM approach seems applicable. So far, the nature of such an accurate prescription of how to behave to get a maximum profit seems unclear.

References

[1] M. Aghassi and D. Bertsimas (2006). Robust game theory. *Mathematical Programming*, B107(1-2), 231–273.

[2] B. Allen and M. Hellwig (1986). Bertrand-Edgeworth oligopoly in large markets. *The Review of Economic Studies*, 53, 175–204.

[3] C. d'Aspremont and J. Gabszewicz (1985). Quasi-Monopolies. *Economica*, 52(206), 141–151.

[4] E. Atiker, W. Neilson, and M. Price (2011) Activity Bias and Focal Points in Extensive Form Games. Working paper, University of Tennessee, Knoxville.

[5] S. Baghestanian and S. Frey (2016). GO Figure: Analytic and Strategic Skills are Separable. *Journal of Behavioral and Experimental Economics.* 64, 71–80.

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

[7] T. Becker, M. Carter, and J. Naeve (2005). Experts playing the traveler's dilemma. Discussion paper 252/2005, Universität Hohenheim.

[8] J. Bertrand, (1883). Review of *Theorie mathematique de la richesse sociale* and of *Recherches sur les principles mathematiques de la theorie des richesses. Journal des Savants*, 67, 499–508.

[9] E. Borel (1953). The Theory of Play and Integral Equations with Skew Symmetric Kernels. *Econometrica*, 21(1), 97–100.

[10] S. Cabrera, M. Capra, and R. Gomez (2007). Behavior in one-shot traveler's dilemma games: model and experiments with advice. *Spanish Economic Review*, 9(2), 129–152.

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

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

[13] S. M. Chowdhury, D. Kovenock, R. M. Sheremeta (2013). An experimental investigation of Colonel Blotto games. *Economic Theory*, 52(3), 833–861.

[14] V.P. Crawford, M.A. Costa-Gomes, N. Iriberri (2013). Structural models of nonequilibrium strategic thinking: theory, evidence, and applications. *Journal of Economic Literature*, 2013, 51, 5–62.

[15] P. Dasgupta and E. Maskin (1986). The existence of equilibrium in discontinuous economic games, II: Applications. *The Review of Economic Studies*, LIII, 27–41.

[16] H.D. Dixon (1984). The existence of mixed-strategy equilibria in a price-setting oligopoly with convex costs. *Economics Letters*, 16, 205–212.

[17] F. Edgeworth, (1925). The pure theory of monopoly (1889), reprinted in *Collected Papers relating to Political Economy 1*. Macmillan.

[18] J.Y. Halpern and R. Pass (2012). Iterated regret minimization: a new solution concept. *Games and Economic Behavior*, 74(1), 194–207.

[19] T. Hayashi (2008). Regret aversion and opportunity dependence. *Journal of Economic Theory*, 139(1), 242–268.

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

[21] K. Modzelewski, J. Stein, J. Yu (2009). An Experimental Study of Classic Colonel Blotto Games. MIT Report 6.207/14.15.

[22] R. Nagel 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.

[23] H.-T. Normann and M. Fonseca (2013). Excess Capacity and Pricing in Bertrand-Edgeworth Markets: Experimental Evidence. *Journal of Institutional and Theoretical Economics*, 169, 199–228.

[24] M. J. Osborne and A. Rubinstein (1994). *A Course in Game Theory.* Cambridge, Mass.: MIT Press.

[25] G. Owen (1968). *Game Theory*. New York: Academic.

[26] L. Renou and K.H. Schlag (2010). Minimax regret and strategic uncertainty. *Journal of Economic Theory*. 145, 264–286.

[27] R.W. Rosenthal (1981). Games of perfect information, predatory pricing, and the chain store paradox. *Journal of Economic Theory* 25(1), 92–100.

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

[29] M. Shubik (1955). A comparison of treatments of a duopoly problem (part II). *Econometrica*, 23, 417–431.

[30] M. Slade (1995) Empirical Games: The Oligopoly Case. *The Canadian Journal of Economics*, 28(2), 368–402.