Research Project Proposal: Regret Minimization for Non-Cooperative Games

Tommaso Bianchi, tommaso4.bianchi@mail.polimi.it

1. Introduction to the problem

1.1. Description of the areas and research topics in which the problem is positioned.

Our research topic is positioned within the area of Algorithmic Game Theory, which lies on the boundary between Mathematics and Computer Science: the former provides the mathematical models to describe the problems and their solutions, in the form of game representations and equilibria concepts; the latter provides the computational and algorithmic tools to either solve those problems in an efficient way or to prove their difficulty[1] to be solved.

More specifically, for Computer Science — the main field of the author — the main research areas involved are: Online Convex Optimization, as most of Game Theoretic problems can be expressed as minimizing convex functions over convex sets; Machine Learning, the area where the concept of regret minimization have originated; Theoretical Computer Science, which is involved each time the complexity or the difficulty [1] of a problem has to be demonstrated.

1.2. Brief description of the research topic.

Suppose a rational agent has to make a decision (say, selecting one out of a set of actions), receiving then back a payoff (or utility) as a function of what he chose, and possibly of what some other agents have chosen. Suppose he has to repeatedly do so, with the objective to get every time the highest possible payoff. Regret minimization tries to approach this problem by defining some strategies (i.e. ways in which the agent can behave, generally represented as probability distribution over the set of actions) that have some particular properties, like for instance the ability to reach the same average payoff as the best constant strategy. This very general settings covers, among others, game theoretical models (i.e. games), thus we can employ regret-based methods to study the convergence to game theoretical equilibria when players repeatedly play the same game.

[1] In terms of computational complexity, e.g., NP-hardness.

1.3. Motivations to support the importance of the research topic.

We think that a regret-based approach for computing game theoretical equilibria is very interesting first and foremost because it conveys the idea of *decentralized rationality*; this means that there is no need of a central device that collects all the data about the game and the players and then compute an equilibrium; instead, rational agents can play the game by themselves correcting their strategy in a way as to minimize their own regret (of whatever kind), and then naturally reach an equilibrium in the long run. Regret minimization is a relevant topic also because it can easily outperform linear programming, the classical way of computing game theoretical equilibria, in particular in very big extensive-form games, paving the way to practical large-scale applications (such as solving Heads-up Hold'em Poker [3]).

1.4. Description of the problem.

In such a context, we aim at solving the following problem: how to efficiently find Coarse Correlated Equilibria (CCE), a particular type of game theoretical solution concept, in games that admit a compact representation, in particular extensive-form and polymatrix ones. Our objective is to produce algorithms able to compute them in general-sum games for both the 2-players case and for the more general n-players one. More precisely, our intent is to produce, for each game representation under evaluation, the following results:

- a regret minimization algorithm able to approach a CCE in polynomial time in the size of the game (already given in the selected compact representation);
- a bound on the growth rate of the regret for each player, tight enough as to show that the average regret goes to zero as the game keep being repeatedly played;
- an evaluation of the quality of the CCE reached by the algorithm, in terms of the so-called Price of Total Anarchy [2];
- an experimental evaluation of the effective performance of the algorithm, both in absolute terms (i.e. how big are the games it can solve in a reasonable amount of time) as well as, when possible, in comparison with other similar algorithms.

1.5. Motivations to support the importance of the problem.

We believe our problem is a relevant one for two main reasons: first, it covers a blind spot of the literature, as — to the best of our knowledge — no one has addressed the problem of CCE computation in extensive-form games, and more in general reseach on regret minimization techniquest for Game Theory tends to concentrate a lot around the problem of computing Nash Equilibria (NE) and refinements of NE; moreover, we think CCE is per se a very promising solution concept, as it is a very simple yet effective method to represent a kind of communication videly used in practice by human beings (i.e. ex-ante correlation, the one typically employed in team card games when communication during the actual match is forbidden, like for instance Bridge) and it possesses very good properties in terms of computational complexity and of attainable social utility.

2. Main related works

The main works related to our research topic are essentially divided between papers dealing with regret minimization in 2-player extensive-form games, with the objective to find a Nash Equilibrium or its various refinements, and papers dealing with more general form of equilibirum — such as CE or CCE — that can be reached by regret-based dynamcs. The first research line is focusing, in the last years, on the Counterfactual Regret Minimization algorithm [8], which allows to compute very precisely approximated equilibria on extensive-form games of considerable size; for the second one, instead, the cornerstone research papers are the ones by Hart and Mas Colell [4, 5], which discuss the existence of simple regret-based procedures approaching, respectively, Correlated and Coarse Correlated Equilibria.

Other important related works include the ones on the so-called Price of Total Anarchy [2], a measure of the quality of a regret-minimizing algorithm in terms of the rate between the optimum social cost and the social welfare effectively achieved; there are also several works related to the use of regret-based approaches for finding equilibria for other game representations, such as Bayesian games [6, 1] or security games [7].

3. Research plan

3.1. Describe the goal of the research.

The goal of our research is to develop algorithms for finding CCE in games with a compact representation — in particular, extensive-form and polymatrix games —, to theoretically prove their soundness and a bound on their average regret and to experimentally test their effectiveness.

3.2. Describe the nature of the research: theory, application, implementation, experimental, hybrid.

Our research is mainly of theoretical nature, as it will focus on finding new algorithmic approaches and to prove their properties. Then, as a side work, we may also develop some prototypal implementations to experimentally test the algorithms in a synthetic scenario.

3.3. Describe the tasks in which the research is decomposed, remarking the output of each task.

Our research can be decomposed in three main tasks — CCE in 2-player extesive-form games, CCE for n-player extensive-form games, CCE for n-players polymatrix games —, each of which can in turn be broken down into four smaller subtasks: development of a regret-based algorithm, proof of a bound on the regret achieved by the algorithm, discussion on the quality of the equilibia reached by the algorithm, referring in particular to the concept of Price of Total Anarchy, and experimental evaluation. For what regards the outputs of each of the main tasks, they will consist mainly of a series of theorems, all with their respective proofs, that set the theoretical grounding of the solutions found, possibly paired with the results coming from the experimental evaluation.

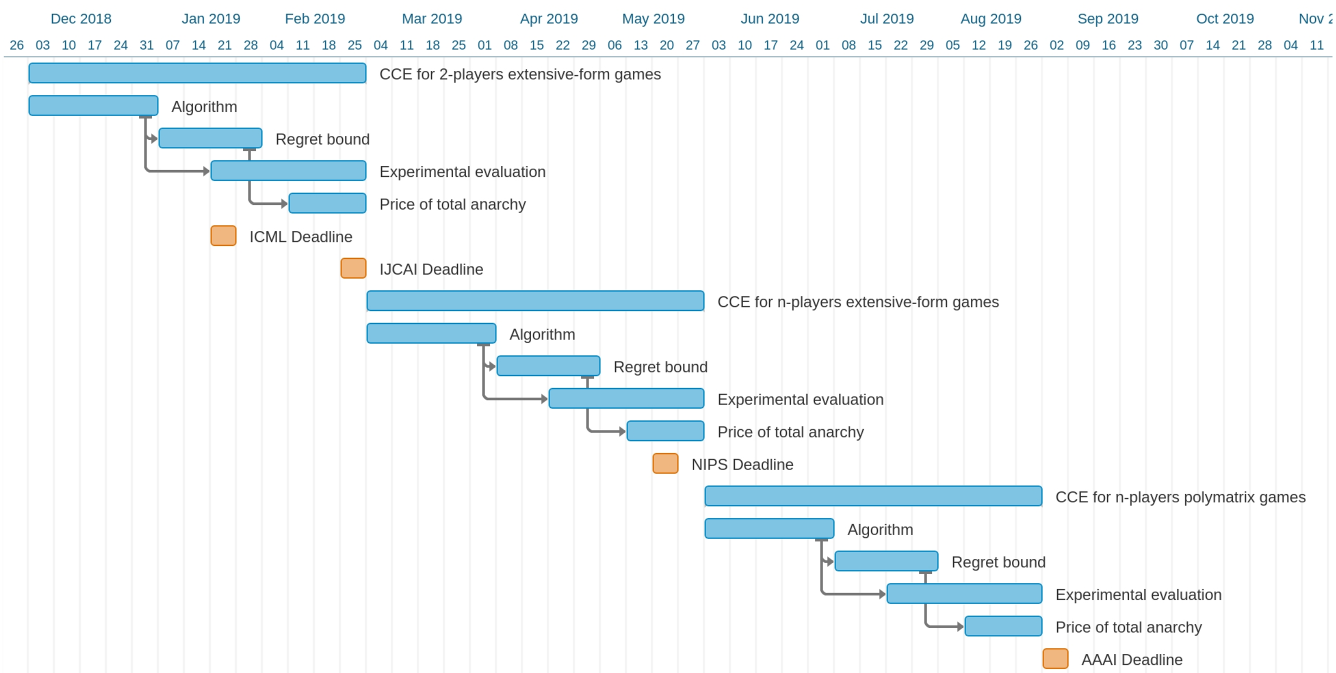

Figure 1: A gantt chart of the tasks of the research project.

3.4. Describe the metrics to use to evaluate the outputs of the research.

The main metrix that will be used to evaluate the output of our research will be:

- the tightness of the regret bound we will be able to achieve, in particular with respect to other analogous ones developed for comparable settings;

- the results of our experimental evaluation, both for the measure of size of the games that we are able to solve as well as for the comparative running time and results quality — in terms of social utility and Price of Total Anarchy —, if other algorithms for the same setting would be available.

References

[1] Balseiro, S., and Gur, Y. Learning in repeated auctions with budgets: Regret minimization and equilibrium.

[2] Blum, A., Hajiaghayi, M., Ligett, K., and Roth, A. Regret minimization and the price of total anarchy. In *Proceedings of the fortieth annual ACM symposium on Theory of computing* (2008), ACM, pp. 373–382.

[3] Brown, N., and Sandholm, T. Superhuman ai for heads-up no-limit poker: Libratus beats top professionals. *Science 359*, 6374 (2018), 418–424.

[4] Hart, S., and Mas-Colell, A. A simple adaptive procedure leading to correlated equilibrium. *Econometrica 68*, 5 (2000), 1127–1150.

[5] Hart, S., and Mas-Colell, A. A general class of adaptive strategies. *Journal of Economic Theory 98*, 1 (2001), 26–54.

[6] Hartline, J., Syrgkanis, V., and Tardos, E. No-regret learning in bayesian games. In *Advances in Neural Information Processing Systems* (2015), pp. 3061–3069.

[7] Lisỳ, V., Davis, T., and Bowling, M. H. Counterfactual regret minimization in sequential security games. In *AAAI* (2016), pp. 544–550.

[8] Zinkevich, M., Johanson, M., Bowling, M., and Piccione, C. Regret minimization in games with incomplete information. In *Advances in neural information processing systems* (2008), pp. 1729–1736.