Research Project Proposal: Function Approximation for Adversarial Team Games

Luca Carminati, luca5.carminati@mail.polimi.it

1. Introduction to the problem

Algorithmic Game Theory is a field of study at the intersection of Game Theory and Computer Science. Its objective is to design algorithms capable of operating strategic decisions in complex environments, optimizing a preference score over the possible outcomes. The complexity of the environment derives from uncertainties due to the presence of imperfect information and/or other agents optimizing their own scores.

1.1. Research topic

During the recent fifteen years, the field of Algorithmic Game Theory has undergone a huge transformation: from being able to solve only small instances of simple zero-sum two-players game, to large complex games such as Poker in case of *Libratus* [4] *Pluribus* [5]. These algorithms can then be used outside their original benchmarks to solve complex real games in the field of security, health and economics.

In particular, the introduction of more and more scalable solving techniques, directly working on the extensive-form representation and making use of sampling, substituted the traditional techniques based on Linear Programming. On top of this, abstraction and function approximation techniques allowed a scale up of the size of solvable games.

In parallel, game theory community shifted part of its interest to other types of games, modeling more complex situations, with three or more players and general-sum payoffs. In these new environments, Nash Equilibria proved fragile in terms of robustness and with poor real-world rational behavior modeling. Thus important research efforts have been made to identify new possible solution concepts and to provide efficient algorithms able to find those solution concepts in reasonable time.

Our research work will be located in the latter research line. Our goal is that of providing a faster and thus more scalable algorithm for one of these equilibrium concepts, by transferring some Function Approximation techniques already employed in the zero-sum case.

1.2. Problem

The specific problem we will address in our research project is the one of finding an equilibrium in Adversarial Team Games, defined in [13]. In class of multiplayer games, players are organized in two teams, and each player has the identical reward as his/her team members, while the game is zero-sum for each outcome.

Adversarial Team Games with two teams are a generalization of two-player zero-sum games, in the sense that if each team is able to perfectly coordinate and cooperate throughout the game, then each team can act as a single player and the game reduces to two-player zero-sum, as shown in [8]. The absence of realtime communication breaks this reduction, and different solution concepts are needed to represensent a rational equilibrium in such contexts. This type of team games are interesting in security and military scenarios, in which communication may not be guaranteed but some forms of coordination are still required to achieve success in the task at hand.

2. Main related works

The main related works in the Adversarial Team Games provide the definitions of the main solution concepts, the bounds on complexity on the general cases, and some algorithms to converge to an equilibrium.

Specifically, Von Stengel and Koller in [13] introduced the concept of Team Maxmin Equilibrium, whereas Basilico et al in [2] analyzed the complexity of finding such an equilibrium in a normal-form game and proposed some algorithms. Celli and Gatti in [8] define the *Team Maxmin Equilibrium (TME)* concept for extensive form games, introducing different variants according to the communication level available for the team members. In particualar, *TMEcor* solution concept is a TME in a game allowing coordination of the team members before the game starts. They also propose an *Hybrid Column Generation* algorithm, iteratively solving a Maxmin and Minmax problem to find the optimal strategy of the team, and leveraging an Integer Linear Program as a Best-Response oracle.

On top of this work, Farina and Celli et al in [9] proposed *Fictitious Team Play*, a generalization of *Fictitious Self Play* [10] applied to team games. In both the works, the crucial bottleneck of computation is represented by the Best-Response computation, currently solved through ILP or Mixed ILP.

Celli et al in [7], Cacciamani et al in [6] developed model-free learning algorithms for team games, leveraging reinforcement learning techniques.

Our research work will develop a model-based algorithm based on [8][9], that addresses the BR computation through an approximated algorithm. This pattern of progressive but theoretically sound approximations will be similar to the one followed for zero-sum games by [3, 10, 11].[1]

3. Research plan

The goal of our research is to provide a more scalable and efficient algorithm to compute a TMEcor in large instances of team games, in the specific case of two players in a team and a single adversary.

In particular, by building on the theoretical framework provided by [9], we are able to transform the original extensive-form team game into an auxiliary two player zero sum game, in such a way that a Nash Equilibrium in the new game is equivalent to a TMEcor in the original game. The challenge is that the player representing the team in the new game has an exponentially large action space. Their proposed solution is that of employing a Fictitious-Play-like learning algorithm, and employing a MILP to compute the Best Response of the team player.

Our intuition is that such a procedure can be made more efficient by employing a CFR-BR [12] algorithm to compute the strategy for the single adversary, while employing an approximated algorithm to compute the BR, opposed to the exact NP-hard MILP formulation used in [9].

For what concerns an approximate algorithm, we propose an Iterative LP formulation similar to [2] as a starting point, but we expect to find a more efficient formulation by employing Multi Agent RL techniques [14]. Possible solutions we are considering also consist in neural optimization frameworks such as in [1].

The nature of our research is thus hybrid: in the first phase we need to theoretically validate our hypotheses, by revising existing literature on the proposed algorithms, while implementing state-of-the art baselines and then incrementally implement our proposed modifications. Thus we expect that the sub tasks constituting our workflow will be:

- **Theoretical guarantees** for the application of CFR-BR in the Auxiliary Game defined in [9]:
- **Best Response problem** formulation and implementation of possible approximating alternatives for the search of a best response:
  - Optimal baseline (OPT):
    - MILP formulation from [8] [9].

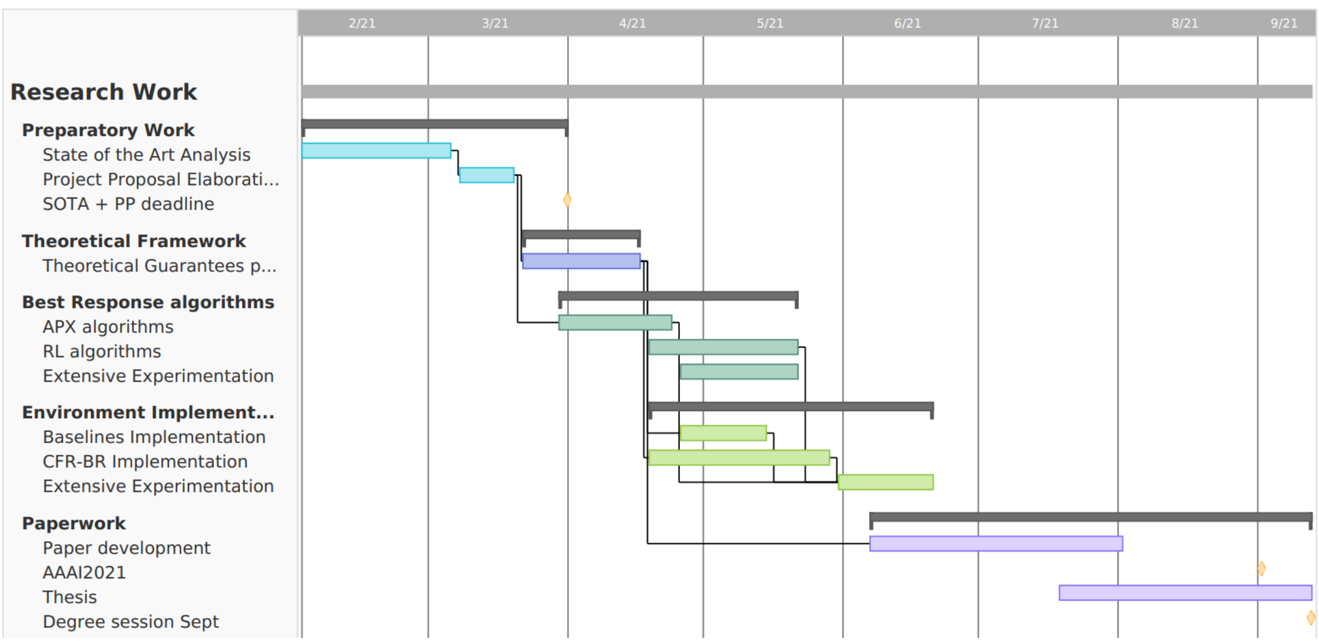
[1] More details on this line of research can be found in the state of the art analysis accompanying this document.

- Approximated iterative (APX):
  * Iterated LP;
  * Monte carlo search.
- Reinforcement Learning approach (RL):
  * Iterative Multi Agent RL algorithms;
  * Deep Multi Agent RL algorithms.

Those will be evaluated according to theirs approximation bounds and time-payoff performances on test games.

- Implementation of our full **framework** and test it against algorithms proposed in [8] [9] in instances of Poker games.

In the following we provide a GANTT diagram of our tasks.



3.1. Research Evaluation

The metrics that will be used to evaluate the output of our research will be based on the following:

- Convergence rate of algorithms;
- Quality of the approximations;
- Performance of experimental implementations, calculated through applications specific results.

All the proposed algorithms will be compared to state-of-the-art algorithms in a unique experimental framework.

References

[1] Bai, S., Kolter, J. Z., and Koltun, V. Deep equilibrium models. *CoRR abs/1909.01377* (2019).

[2] Basilico, N., Celli, A., Nittis, G. D., and Gatti, N. Team-maxmin equilibrium: efficiency bounds and algorithms, 2016.

[3] Brown, N., Lerer, A., Gross, S., and Sandholm, T. Deep Counterfactual Regret Minimization. *arXiv:1811.00164 [cs]* (May 2019). arXiv: 1811.00164.

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

[5] Brown, N., and Sandholm, T. Superhuman ai for multiplayer poker. *Science 365*, 6456 (2019), 885–890.

[6] Cacciamani, F., Celli, A., Ciccone, M., and Gatti, N. Multi-agent coordination in adversarial environments through signal mediated strategies. *arXiv:2102.05026 [cs]* (Feb 2021). arXiv: 2102.05026.

[7] Celli, A., Ciccone, M., Bongo, R., and Gatti, N. Coordination in adversarial sequential team games via multi-agent deep reinforcement learning. *CoRR abs/1912.07712* (2019).

[8] Celli, A., and Gatti, N. Computational results for extensive-form adversarial team games. *arXiv:1711.06930 [cs]* (Nov 2017). arXiv: 1711.06930.

[9] Farina, G., Gatti, N., Celli, A., and Sandholm, T. Ex ante coordination and collusion in zero-sum multi-player extensive-form games. 17.

[10] Heinrich, J., and Lanctot, M. Fictitious Self-Play in Extensive-Form Games. 9.

[11] Heinrich, J., and Silver, D. Deep Reinforcement Learning from Self-Play in Imperfect-Information Games. *arXiv:1603.01121 [cs]* (June 2016). arXiv: 1603.01121.

[12] Johanson, M., Bard, N., Burch, N., and Bowling, M. Finding Optimal Abstract Strategies in Extensive-Form Games. 9.

[13] von Stengel, B., and Koller, D. Team-maxmin equilibria. *Games and Economic Behavior 21*, 1 (1997), 309–321.

[14] Zhang, K., Yang, Z., and Başar, T. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *arXiv:1911.10635 [cs, stat]* (Nov 2019). arXiv: 1911.10635.