State of the Art on: Regret Minimization for Non-Cooperative Games

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

1. Introduction to the research topic

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.

Being the research topic at the intersection between different disciplines, the most relevant venues for publishing scientific papers are also spread between different research areas, with a particular enphasis on Artificial Intelligence — which also comprises Machine Learning — as the work of the author is more pertaining to it. They have been selected on the basis of these factors: measured quality of the venue, represented by the GGS Ranking[2] for conference and by the impact factor for journals, acceptance rate, publication of the most influential articles and authors in their respective areas and opinion of the scholars working in these areas, in particular the ones the author is collaborating with.

The most prestigious conferences related to Regret Minimization for Non-Cooperative Games, with the research area they belong to, are:

- Association for the Advancement of Artificial Intelligence (AAAI) - Artificial Intelligence;
- International Conference on Machine Learning (ICML) - Artificial Intelligence;
- International Joint Conference on Artificial Intelligence (IJCAI) - Artificial Intelligence;
- Neural Information Processing Systems (NIPS) - Artificial Intelligence;

[1] In terms of computational complexity, e.g., NP-hardness.
[2] See the Website: `http://gii-grin-scie-rating.scie.es/conferenceRating.jsf`.

- ACM Symposium on Theory of Computing (STOC) - Theoretical Computer Science;
- Autonomous Agents and Multi Agent Systems (AAMAS) - Artificial Intelligence;
- ACM Conference on Economics and Computation (EC) - Game Theory, Microeconomics, Artificial Intelligence, Theoretical Computer Science;
- Association for Uncertainty in Artificial Intelligence (UAI) - Artificial Intelligence;
- ACM-SIAM Symposium on Discrete Algorithms (SODA) - Theoretical Computer Science.

The most prestigious journals related to Regret Minimization for Non-Cooperative Games, with the research area they belong to, are:

- Artificial Intelligence Journal - Artificial Intelligence;
- Journal of Artificial Intelligence Research - Artificial Intelligence;
- Machine Learning - Artificial Intelligence;
- Journal of Machine Learning Research - Artificial Intelligence;
- Algorithmica - Theoretical Computer Science;
- Games and Economic Behavior - Game Theory, Economics.

1.1. Preliminaries

We briefly introduce several of the basic concepts needed to understand what Game Theory is about and then to be able to frame and classify the main works that have been done in this field. Further details can be found in [29].

1.1.1 Game Representations

Definition 1. *A* normal-form *(or* strategic-form*) game is a tuple* (N, A, U)*, where:*

- N *is the set of players;*
- $A = \times_{i \in N} A_i$ *is the set of action profiles (i.e., tuples containing an action for each player), where* A_i *is the set of actions of player* i*;*
- $U = (U_1, \ldots, U_n)$ *is the set of the utility functions* $U_i : A \to \mathbb{R}$*, each mapping an action profile into its respective payoff for player* i*.*

Definition 2. *An* imperfect-information extensive-form game Γ *is a tuple* (N, V, H, A, L, χ, U) *where:*

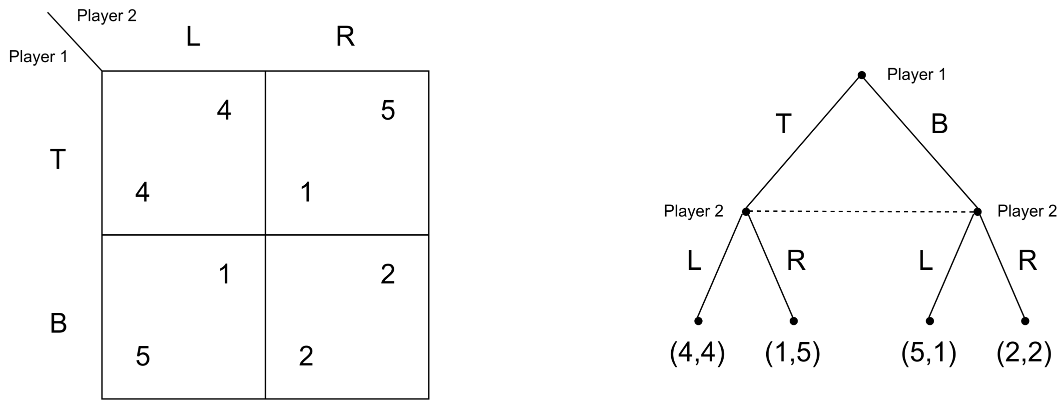

Figure 1: A two-player normal-form game and its equivalent extensive-form game.

- *N is the set of n players;*
- *V is the set of nonterminal decision nodes, and $V_i \subseteq V$ is the set of decision nodes belonging to player $i \in N$;*
- *$H = (H_1, \dots, H_n)$ is the collection of information sets; for each $i \in N$, H_i is an information partition of V_i such that decision nodes within the same information set $h \in H_i$ are not distinguishable by player i;*
- *$A = \times_{h \in H} C_h$ is the set of actions, where C_h is the set of actions available at information set h (w.l.g. we assume each C_h disjoint);*
- *L is the set of terminal nodes (leaves);*
- *$\chi : V \times A \to V \cup L$ is the successor function;*
- *$U = (U_1, \dots, U_n)$ is the set of utility functions $U_i : L \to \mathbb{R}$, $\forall i \in N$.*

Moreover, an extensive-form game where, at each stage, the players recall the whole information they acquired in earlier stages is said to have *perfect recall*, while if it lacks this property it is said to have *imperfect recall*. More formally, if for all $h \in H_i$ and $n \in h$ we call $\rho(n) \subset H_i$ the set of information sets of player i that are on the path from the root of the tree to n, then player i has perfect recall if and only if $\rho(n_1) = \rho(n_2)$ for all $h \in H_i$ and $n_1, n_2 \in h$.

An extensive-form game Γ can be equivalently represented in normal-form. Let $P_i = \times_{h \in H_i} \rho(h)$ be the set of pure normal-form plans of player $i \in N$. A normal-form plan $p \in P_i$ specifies an action per information set of player i. The normal-form game equivalent to $\Gamma = (N, A, V, L, \iota, \rho, \chi, U, H)$ is then the game $(N, P = \times_{i \in N} P_i, U')$, where $U' = (U'_1, \dots, U'_n)$ is the set of the $U'_i : P \to \mathbb{R}$ s.t. $U'_i(p_1, \dots, p_n) = U_i(\ell)$, where $\ell \in L$ is the terminal node reached when playing plan profile $(p_1, \dots, p_n)$. The size of this the normal-form game is, in general, exponential in the size of Γ.

A game, either in normal-form or in extensive-form, is said to be a zero-sum game if

$\sum_{i \in N} U_i(x) = 0 \; \forall x \in Dom(U)$; otherwise, the game is said to be a general-sum game.

1.1.2 Strategy Representations

A normal-form strategy s_i for $i \in N$ is defined as a function $s_i : A_i \rightarrow \Delta^{|A_i|}$. We denote by S_i the normal-form strategy space of player i, and by $S = \times_{i \in N} S_i$ the space of normal-form strategy profiles.

For extensive-form games, as normal-form strategies are exponentially many in the size of the game, a concept often used is the one of behavioural strategies [21]. A behavioural strategy σ_i for $i \in N$ is defined (with a slight abuse of notation) as a function $\sigma_i : H_i \rightarrow \Delta^{|C_{H_i}|}$, that assigns to each information set $h \in H_i$ of player i a distribution over its actions C_h. We denote by Σ_i the behavioural strategy space of player i, and by $\Sigma = \times_{i \in N} \Sigma_i$ the space of behavioural strategy profiles.

1.1.3 Equilibria

There are three classical notions of equilibria for normal-form games: Nash Equilibrium [27], Correlated Equilibrium [3] and Coarse Correlated Equilibrium [26].

For convenience of notation, let $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+i}, \ldots, s_n) \in \times_{j \in N \setminus \{i\}} S_j$ be the strategy profile s without the strategy for player i, and let $U_i(s) = \sum_{a \in A} (\prod_{s_p \in s} s_p(a)) U_i(a)$ be the expected utility for player i under strategy profile s.

Definition 3. *Given a normal form game (N, P, U), a strategy profile $\hat{s} \in S$ is a Nash Equilibrium (NE) if and only if, for every $i \in N$, the following holds:*

$$U_i(\hat{s}) \geq \max_{s_i \in S_i} U_i(s_i, \hat{s}_{-i}).$$

A *correlated* (joint) normal-form strategy $x \in X$ is defined as $x : A \rightarrow \Delta^{|A|}$. It is worth noting that, in general, a joint strategy cannot be marginalized into a strategy profile s such that $\prod_{s_p \in s} s_p(a) = x(a)$ for each action $a \in A$.

Definition 4. *Given a normal form game (N, P, U), a joint strategy $x \in X$ is a Correlated Equilibrium (CE) if and only if, for every $i \in N$ and $a_i, a_i' \in A_i$, the following holds:*

$$\sum_{a_{-i} \in A_{-i}} x(a_i, a_{-i})(U(a_i, a_{-i}) - U(a_i', a_{-i})) \geq 0.$$

A CE can be interpreted in terms of a mediator (a.k.a. correlation device) who, *ex ante* the play, draws $(a_1, \ldots, a_n)$ according to the publicly known x and privately communicates each *recommendation* a_i to the corresponding player i.

Definition 5. *Given a normal form game* (N, P, U)*, a joint strategy* $x \in X$ *is a Coarse Correlated Equilibrium (CCE) if and only if, for every* $i \in N$ *and* $a'_i \in A_i$*, the following holds:*

$$\sum_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} x(a_i, a_{-i})(U(a_i, a_{-i}) - U(a'_i, a_{-i})) \geq 0.$$

CCE differs from CE in that a CCE only requires that following the suggested action is a best response in expectation, before the recommended action is actually revealed Moreover, we recall the inclusion relation between all these equilibria, which is, in general, strict: every NE is also a CE, but a CE may be not a NE, and every CE is also a CCE, but a CCE may be not a CE.

All of these equilibrium concepts can be extended to extensive-form games as well, by leveraging the equivalent normal-form representation; in general, this leads to exponentially complex problems due to the rise in dimensionality required by this equivalence, but in some particular cases (e.g. two-player zero-sum) it can be shown that alternative techniques can be employed to find extensive-form equilibria without the need of the exponentially-large normal-form.

1.1.4 Regret and Hannan Consistency

The notion of regret refers to a very different setting with respect to the "static" equations for equilibria that we have seen so far: the game is repeatedly played, with a potentially different strategy profile $s^t \in S$ at each timestep t, and players can at each timestep observe their own payoff and the payoff the would have received for each action they could have played (in Machine Learning such a setting is usually called *expert*). Regret is, in general, a measure of how much a player would have preferred to play a different strategy with respect to s^t, the strategy he actually used. As regret-based algorithms leverage past information (the so-called history of play) to learn how to play a game — possibly reaching an equilibrium —, they are called learning algorithms (or sometimes forecasters).

The classical concept of *Hannan Consistency* [15] (also known as *Universal Consistency*) is the ability of a learning algorithm to achieve a sublinear regret with respect to the best possible action in hindsight. More formally:

Definition 6. *A learning algorithm* Ł *is said to be Hannan Consistent for player* $i \in N$ *if and only if*

$$\lim_{T \to \infty} \sup \frac{1}{T} \left(\max_{a_i \in A_i} \sum_{t=1}^{T} U_i(a_i, s^t_{-i}) - \sum_{t=1}^{T} U_i(s^t) \right) = 0.$$

The regret minimized by Hannan Consistency, known as *external regret*, is only one of the possible ways to define regret; others include *internal regret*, that is computed over pairs of actions, and *swap regret*, where arbitrary substitutions of actions are taken into account.

1.1.5 Approachability

The concept of approachability of a convex set, introduced by Blackwell in 1956 [5], is a fundamental tool for the analisys of dynamic regret-based algorithms. In its more general setting, it involves two players, called i and $-i$, repeatedly playing a vector-valued game (i.e. a game where the utility function $\vec{U}$ goes from the action space to some vector space $\mathbb{R}^L$) in which player i tries to keep the average payoff as close as possible to a given convex closed set $\mathscr{C}$, while player $-i$ tries to deny that.

Theorem 1 (Blackwell's Approachability Theorem)**.** *Let $\mathscr{C} \in \mathbb{R}^L$ be a convex and closed set, and denote by $w_{\mathscr{C}}$ its support function (i.e. $w_{\mathscr{C}}(\lambda) = sup\{\lambda \cdot c : c \in \mathscr{C}\}$ for all $\lambda \in \mathbb{R}^L$). Then $\mathscr{C}$ is approachable by player i if and only if for every $\lambda \in \mathbb{R}^L$ there exists a mixed strategy $s_\lambda \in \Delta(A_i)$ such that, for all $a_{-i} \in A_{-i}$*

$$\lambda \cdot \vec{U}(s_\lambda, a_{-i}) \leq w_{\mathscr{C}}(\lambda).$$

In the context of regret minimization, this result is often used to show that a procedure is able to minimize regret, by substituting the negative orthant $\mathbb{R}^L_- = \{x \in \mathbb{R}^L : x \leq 0\}$ as the set $\mathscr{C}$ to approach and by setting the vector payoff of the game to the vector of regret observed by player i at timestep t.

1.2. Research topic

The main objective of our research topic is to employ regret minimization techniques to approach game theoretical equilibria in a computationally efficient way. While doing so, we also aim at studying some important properties of both the game itself and its equilibria, such as the possibility to compactly represent them (i.e. store games and their equilibria in an efficient way). More specifically, our research will focus on Correlated and Coarse Correlated Equilibria for games in extensive-form, as the available literature lacks strong results for this very interesting setting.

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 an interesting topic also because it can easily outperform linear programming, the classical way of computing game theoretical equilibria, in particular on extensive-form games, where often the problem can be formulated as a set of local regret minimizers at each information set, thus fitting in a more natural way the structure of the game tree. Another important advantage over linear programming, in particular for practical large-scale applications (such as solving Heads-up

Hold'em Poker [7]), is the fact that regret-based algorithms are almost always approximate anytime algorithms, which means that their computation can be stopped at any moment and still yield an appoximate solution (of course, the more time is given to an algorithm the better the approximation).

As regards the choice of CE and CCE as solution concepts in general-sum games, it is motivated by these facts: they are generalization of NEs, thus they are in a sense richer sets, which include equilibria with better properties in terms of social welfare (i.e. the sum of the utility for all the players) — a CE may lead to a social welfare arbitrarily larger than a NE and a CCE may lead to a social welfare arbitrarily better than a CE; they are computationally more tractable than NEs, which makes the problem more meaningful in an algorithmic perspective; their formulation is mathematically more similar to some forms of regret, which makes more natural to employ regret-based techniques in a very general way (while for NE it is possible to adapt regret minimizers only for 2-player zero-sum games) — it is known that regret dynamics may lead to CCEs that are close in terms of social welfare to the socially optimal CCE.

We are also focusing on games that admit a much more compact representations than their equivalent normal-form ones, and in particular on extensive-form games — far more expressive than normal-form games while still being much more general than most of the other, more specialized, game representations — and on polymatrix games.

2. Main related works

2.1. Classification of the main related works

The most relevant dimensions alongh which we can classify the available literature about regret minimization for non-cooperative game theory are the following.

- The kind of equilibrium concept is reached (more formally, approached) by a particular regret-based algorithm, taking into consideration NEs, CEs, and CCEs.
- The number of players of the game (either two or a general number N).
- The type of game being played, in particular zero-sum or general-sum.
- The game representation, in particular normal-form or extensive-form.

	2 players zero-sum	**N players zero-sum**	**N players general-sum**
NE normal-form	[1], [2]		[10], [11]
NE extensive-form	[8], [9], [20], [22][3], [23], [30], [31]	[13], [18], [28]	[12]
CE normal-form	[16]		
CE extensive-form			
CCE normal-form	[17]		
CCE bayesian	[19]		
CCE extensive-form			
Other	[4], [6], [14], [24], [25]		

2.2. Brief description of the main related works

The state of the art regret-based algorithm for finding 2-player zero-sum NE in normal-form games is the EXP3 algorithm [2], which can also be extended to compute CE and CCE: unfortunately, in this case regret is no longer guaranteed to be minimized, and so this approach become ineffective. In principle, this result could also be applied to computing NE in extensive-form games, but in practice it would suffer from the exponentially increased dimensionality of the equivalent normal-form game. For this reason, research has searched for more specialized algorithms able to exploit the regularities in the structure of compactedly represented games (like extensive-form games) and to thus be able to efficiently compute equilibria. Another important family of algorithms employed to compute NE in 2-player zero-sum games efficiently are the so-called Multiplicative Weights (or Polynomial Weights) algorithms [1], that can be proved have optimal worst-case expected regret (i.e. it matches best achievable regret lower bound [4] up to constant factors). There are also some works [10, 11] tackling the problem of NE in normal-form general-sum games with an arbitrary number of player, although they are still heavily theoretical in their nature and there has not been yet any algorithmic approach on this research line.

As it is evident from the table, most of the research so far has been focused on finding 2-player zero-sum NE in extensive-form games (with perfect recall), and in particular has been concentrated on the Counterfactual Regret Minimization (CFR) algorithm [5] [31] and its many variants and adjustments, such as CFR+ [30] and Monte Carlo CFR [23, 20]; a CFR-like algorithm has also been used to compute perfect equilibria [8], which are NE robust to players' deviations from the equilibrium itself, and quantal response equilibria [9], where players are

[3] In this work, players may have a specific form of imperfect recallness.

[4] Any regret minimizing algorithm have a $\Omega(\sqrt{Tln|A|})$ regret.

[5] In its first version, CFR had a regret upper bound for each player of $O(\Delta_{u,i}|I_i|\sqrt{T|A_i|})$, where $\Delta_{u,i}$ is the range of the utility of player i.

assumed to make errors in choosing which pure strategy to play. Some works also tried to tackle the problem of extending the CFR algorithm to the N-player setting [28, 13] or to the general-sum one [12], even though there is still no theoretical guarantee of convergence to a NE for this cases. Another proposed extension, also without solid theoretical guarantess, is the one by [18], focusing on the case of multiplayer collaborative games (2 teams of 2 players). There is then only one work that tries to tackle imperfect-recall extensive-form games [22], achieving some theoretical results only on a very restricted class of games (the so-called well-formed imperfect-recall games). This wide interest is Counterfactual Regret Minimization as a mean for reaching NE, particularly in large-scale extensive-form games, is justified by its incredibly good scalability properties; as a lot of experimental work suggest, CFR is by far the fastest algorithm to date for computing 2-player zero-sum NE.

For what regards CE and CCE, the main works are the one by Hart and Mas Colell [16, 17], that ground the main equivalence between minimizing regret, respectively internal or external regret, to approach them in normal-form general-sum games. To the best of our knowledge, there has been no effort still to generalize those results to the extensive-form setting.

There is a duality in the form of regret that all of this papers focus on, which are sometimes called *action regret* and *distribution regret*: the former is calculated with respect to the actions that the agent actually plays, which is in general sampled from a distribution over the action set (i.e. a strategy); in the latter case, instead, the regret at time t is directly calculated with respect to the agent's mixed strategy itself — essentially, it is an expectation at time $t-1$ of the action regret at time t. In particular, most of the works done on CFR employ distribution regret (at least from the theoretical point of view), while the literature on CE and CCE tends more towards employing action regret.

Among other works that do not precisely fit in the classification we have proposed, we can find both very theoretical results indipendent of the specific equilibrium concept reached, such as the Price of Total Anarchy [6, 25] — a concept related to the quality of the equilibirum reached —, and more practical results that apply to specific class of games, such as Bayesian Games [19, 4], security games [24] or Convex Games [14].

2.3. Discussion

To conclude our review of the available literature on Regret Minimization for Non-Cooperative Games, we present a critical examination of what have been the focal points of the research in the last years, underlining what problems and issues are still open and where are the areas in which further work is needed.

The largest amount of contributions has certainly and clearly been the ones targeting 2-player zero-sum Nash Equilibria computation, with a particular focus on alternative game representations with respect to the normal-form (e.g. extensive-form games, bayesian games);

some of the techniques developed in this area have proved to be very efficient and scalable, enabling for the first time to solve some really big games, including complex ones competitively played by humans (e.g. Heads-up Hold'em Poker [7]).

Although it has paved the way for such remarkable achievements, this strong committment to a single branch of the research field has left a lot of other equally interesting ones without adequate attention. For instance, not a lot is known about what kind of equilibria can be efficiently reached in imperfect-recall games, nor any algorithm has been developed to do so. The same holds for the computation of CE and CCE (and in general of communication-related equilibria) for compactly represented games, whether they be extensive-form, polymatrix ones or others. We believe that a lot more can be done in these areas, with the help of the powerful tool of regret-based learning.

References

[1] Arora, S., Hazan, E., and Kale, S. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing 8*, 1 (2012), 121–164.

[2] Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *FOCS* (1995), p. 322.

[3] Aumann, R. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics 1*, 1 (1974), 67–96.

[4] Balseiro, S., and Gur, Y. Learning in repeated auctions with budgets: Regret minimization and equilibrium. In *EC* (2017).

[5] Blackwell, D., et al. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics 6*, 1 (1956), 1–8.

[6] Blum, A., Hajiaghayi, M., Ligett, K., and Roth, A. Regret minimization and the price of total anarchy. In *STOC* (2008), pp. 373–382.

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

[8] Farina, G., Kroer, C., and Sandholm, T. Regret minimization in behaviorally-constrained zero-sum games. *arXiv preprint arXiv:1711.03441* (2017).

[9] Farina, G., Kroer, C., and Sandholm, T. Online convex optimization for sequential decision processes and extensive-form games. *arXiv preprint arXiv:1809.03075* (2018).

[10] Foster, D. P., and Young, H. P. Regret testing: learning to play nash equilibrium without knowing you have an opponent. *Theoretical Economics 1*, 3 (2006), 341–367.

[11] Germano, F., and Lugosi, G. Global nash convergence of foster and young's regret testing. *Games and Economic Behavior 60*, 1 (2007), 135–154.

[12] Gibson, R. Regret minimization in non-zero-sum games with applications to building champion multiplayer computer poker agents. *arXiv preprint arXiv:1305.0034* (2013).

[13] Gibson, R., and Szafron, D. Regret minimization in multiplayer extensive games. In *IJCAI* (2011), vol. 22, p. 2802.

[14] Gordon, G. J., Greenwald, A., and Marks, C. No-regret learning in convex games. In *ICML* (2008), pp. 360–367.

[15] Hannan, J. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games 3* (1957), 97–139.

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

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

[18] Hartley, M., Zheng, S., and Yue, Y. Multi-agent counterfactual regret minimization for partial-information collaborative games. *White paper at CMU*.

[19] Hartline, J., Syrgkanis, V., and Tardos, E. No-regret learning in bayesian games. In *NIPS* (2015), pp. 3061–3069.

[20] Jackson, E. Slumbot nl: Solving large games with counterfactual regret minimization using sampling and distributed processing. In *AAAI Workshop on Computer Poker and Incomplete Information* (2013).

[21] Kuhn, H. Extensive games and the problem of information, contributions to the theory of games ii, kuhn, hw and aw tucker, eds. 193-216, 1953.

[22] Lanctot, M., Gibson, R., Burch, N., Zinkevich, M., and Bowling, M. No-regret learning in extensive-form games with imperfect recall. *arXiv preprint arXiv:1205.0622* (2012).

[23] Lanctot, M., Waugh, K., Zinkevich, M., and Bowling, M. Monte carlo sampling for regret minimization in extensive games. In *NIPS* (2009), pp. 1078–1086.

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

[25] Monnot, B., and Piliouras, G. Limits and limitations of no-regret learning in games. *The Knowledge Engineering Review 32* (2017).

[26] Moulin, H., and Vial, J.-P. Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory 7*, 3 (1978), 201–221.

[27] Nash, J. Non-cooperative games. *Annals of mathematics* (1951), 286–295.

[28] Risk, N. A., and Szafron, D. Using counterfactual regret minimization to create competitive multiplayer poker agents. In *AAMAS* (2010), pp. 159–166.

[29] Shoham, Y., and Leyton-Brown, K. Multiagent systems: Algorithmic, game-theoretic, and logical foundations, 2009.

[30] Tammelin, O. Solving large imperfect information games using cfr+. *arXiv preprint arXiv:1407.5042* (2014).

[31] Zinkevich, M., Johanson, M., Bowling, M., and Piccione, C. Regret minimization in games with incomplete information. In *NIPS* (2008), pp. 1729–1736.