Anchoring Theory in Sequential Stackelberg Games

Extended Abstract

Jan Karwowski
Warsaw University of Technology
Warsaw, Poland
jan.karwowski@mini.pw.edu.pl

Jacek Mańdziuk
Warsaw University of Technology
Warsaw, Poland
j.mandziuk@mini.pw.edu.pl

Adam Żychowski
Warsaw University of Technology
Warsaw, Poland
a.zychowski@mini.pw.edu.pl

ABSTRACT

An underlying assumption of Stackelberg Games (SGs) is perfect rationality of the players. However, in real-life situations the followers (thieves, poachers, smugglers), as humans in general, may act not in a perfectly rational way, since their decisions may be affected by biases of various kinds which bound rationality of their decisions. One of the popular models of bounded rationality is Anchoring Theory (AT) which claims that humans have a tendency to flatten probabilities of available options, i.e. they perceive a distribution of these probabilities as being closer to the uniform distribution than it really is. We propose an efficient formulation of AT in sequential extensive-form SGs suitable for Mixed-Integer Linear Program solution methods and compare the results of its implementation in five state-of-the-art methods for solving sequential SGs.

KEYWORDS

Bounded rationality; Anchoring Theory; Stackelberg Games

ACM Reference Format:
Jan Karwowski, Jacek Mańdziuk, and Adam Żychowski. 2020. Anchoring Theory in Sequential Stackelberg Games. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

Bounded rationality (BR) [15] in problem-solving refers to limitations of decision-makers that lead them to taking non-optimal actions. Except for limited cognitive abilities, BR can be attributed to partial knowledge about the problem, limited resources, or imprecisely defined goal [1, 14]. The most popular models of BR are Prospect Theory [5], Anchoring Theory (AT) [16], Quantal Response [11] and Framing Effect [17]. In this paper, AT approach implemented in COBRA [12, 13] for normal-form games is extended to the case of *sequential extensive-form games* in a way that avoids non-linear constraints, which makes it suitable for a wide range of MILP approaches. AT assumes the existence of a certain distortion (towards the uniform distribution of probabilities of possible actions) of the follower's perception of the leader's mixed strategy. The leader being aware of that distortion can exploit this weakness in their strategy formulation.

A *pure strategy* of the player is an assignment of one action to each *potentially reachable* state of the game. Let's denote a set of all pure strategies of player i by Π_i. A *mixed strategy* δ_i is a probability distribution over Π_i. In *extensive-form* SGs each node in a game tree is uniquely defined by a pair of sequences: the leader's actions and the follower's actions which lead to that node. These sequences will be denoted by σ_l and σ_f, resp.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

The goal of SG is to find Strong Stackelberg Equilibrium [9] i.e. a strategy profile (δ_l^*, δ_f^*) satisfying the two following equations: $\delta_l^* = \arg\max_{\delta_l} u_l(\delta_l, \delta_f^*)$ and $\delta_f^* = \arg\max_{\delta_f} u_f(\delta_l, \delta_f)$. The second one defines the follower's best (optimal) response to the leader's strategy δ_l while the first one selects the best leaders's strategy against the optimal follower's response. Additionally it is assumed that the follower breaks ties in favour of the leader. $u_i, i \in \{l, f\}$ is the *utility/payoff* of player i.

2 AT IN SEQUENTIAL SG (ATSG)

ATSG is implemented as a distorted follower's perception of the leader's behavior strategy. Let's denote by $q(i)$ a probability of choosing action i by the leader in a given *information set* (IS), stemming from its behavior strategy. The most straightforward implementation of AT (though non-linear in sequence-form games) is to change the probability of taking this action to $q'(i) = (1 - \alpha q(i)) + \alpha/M$, where M is the number of actions available in this IS. However, in sequence-form games, for a given leader's feasible sequence of actions $\sigma_l = a_1, a_2, a_3, \ldots, a_n$ a probability of playing it, based on behavior strategy, would be $p(\sigma_l) = q(a_1)q(a_2)\cdots q(a_n)$ and the distorted AT probability would become (*): $p'(\sigma_l) = ((1-\alpha)q(a_1) + \alpha/M_1)((1-\alpha)q(a_2) + \alpha/M_2)\cdots((1-\alpha)q(a_n) + \alpha/M_n)$, where M_i is the number of actions available in IS in which a_i is played.

State of the art approaches to SSE in extensive-form games utilize MILP formulations capable of exploiting a sequence form of a game [2, 3]. Variables p in MILP formulation of SG are products of $q(a_i)$ values presented above (*), and as such cannot be expressed in a linear form with respect to $q(a_i)$. Consequently, applying the above AT modification to MILPwould end-up with non-linear constraints, inadequate for MILP formulation. Consequently, we propose to simplify the above ATSG by dropping distortion coefficients from all but the last one probabilities (**): $p''(\sigma_l) = q(a_1)\cdots q(a_{n-1})((1-\alpha)q_{a_n} + \alpha/M_n) = q(a_1)\cdots q(a_{n-1})\cdot \alpha/M_n + (1-\alpha)q(a_1)\cdots q(a_{n-1})q(a_n) = p(init(\sigma_l))\alpha/M_n + (1-\alpha)p(\sigma_l)$, where $init(\cdot)$ is a function which outputs a sequence without the last move. A simplified version of ATSG (**) is well suited to MILP/LP formulations of sequence-form games.

Please note that relations among probabilities of the leader's actions within a single IS are the same according to both equations (*) and (**), i.e. $\forall \sigma_l^1, \sigma_l^2 \quad I(\sigma_l^1) = I(\sigma_l^2) \Rightarrow p'(\sigma_l^1)/p'(\sigma_l^2) = p''(\sigma_l^1)/p''(\sigma_l^2)$, where $p'(\sigma), p''(\sigma)$ represent probability of sequence σ in a given IS calculated according to (*) and (**), resp. Furthermore, for a given sequence σ_l, for small values of α a difference $|p''(\sigma_l) - p'(\sigma_l)|$ is also small.

Modification of MILP/LP based methods. ATSG formulation (**) was incorporated into three state-of-the-art MILP methods for general-sum sequential SGs: *BC2015* [2], *C2016* [3] and *CBK2018* [4]. The first two are exact methods, the last one provides approximate solutions. In each case, due to inherent requirement of constraints linearity, a simplified AT version (**) was used, i.e. each occurrence of $p(\sigma_l)$ in constraints enforcing the followers optimal response (either stand-alone or as a joint probability $p(\sigma_l, \sigma_f)$) in the original MILP formulations of these methods was replaced by $p(init(\sigma_l))\alpha/M_n + (1-\alpha)p(\sigma_l)$.

Modification of heuristic methods. The remaining two methods are heuristic non-MILP approaches to solving sequential extensive-form SGs: *O2UCT* [7, 8] and *EASG* [18]. The former (double-oracle UCT sampling) relies on a guided sampling of the follower's strategy space interleaved with finding a feasible leader's strategy using double-oracle method. The latter utilizes Evolutionary Algorithm (EA) to find the leader's mixed strategy.

ATSG implementation in *O2UCT* required using distorted probabilities (**) in the follower's oracle when calculating the expected value, as well as in a procedure that calculates a difference between the follower's utilities for two strategies.

Incorporation of ATSG into *EASG* relies on considering a distorted version (**) of the leader's mixed strategy when calculating the best follower's response against which each chromosome is evaluated.

Observe that *O2UCT* and *EASG* are flexible in adoption of various ATSG formulations. For both methods, contrary to MILP/LP ATSG implementations, the potential existence of non-linearities in the formulas defining distorted follower's probabilities is not harmful, and - in principle - any other BR modification could be used instead of eq. (**). For comparability reasons, we will use a linear form (**) in the experiments.

3 EXPERIMENTAL EVALUATION

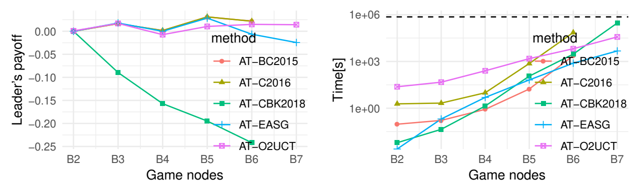
In what follows modifications of considered methods incorporating ATSG will be referred to with the prefix *AT-*.

Benchmark games. Experimental evaluation was performed on a set of patrolling *Warehouse Games* introduced in [6]. Game instances can be downloaded from our project website [10]. The benchmark set consisted of 25 games generated on 4×4 grid, $T = 3, \ldots, 7$, albeit for $T = 7$ exact methods were unable to compute solutions with allotted time and memory.

Experimental setup. For each game instance (game layout and game length) *AT-O2UCT* and *AT-EASG* were run 10 times and for each other (deterministic) MILP method a single trial was performed. Tests were run on Intel Xeon Silver 4116 @ 2.10GHz with 256GB RAM. Experiments with *AT-O2UCT* and *AT-EASG* were run in parallel, each with 8GB RAM assigned. Tests with *AT-C2016*, *AT-CBK2018*, *AT-BC2015* were run sequentially with all 256GB RAM available in each trial. All tests were limited to 200 hours (per single test) and forcibly terminated if not completed within the allotted time. Results for all games are presented w.r.t the aggregated number of game nodes ($|\mathcal{H}|$) in the extensive-form game representation. This grouping followed the formula: $bucket = 10^{round(\log_{10}|\mathcal{H}|)}$, where *round* rounds a number to the nearest integer. Henceforth $B_i, i = 2, \ldots, 7$ will denote the *i-th bucket of games*.

Payoffs. The average expected leader's utilities are compared in Fig. 1a. *AT-C2016* and *AT-BC2015* are exact methods, so their results are clearly the highest and the resp. plots overlap. Both non-MILP heuristic methods perform slightly worse, although for games from $B_{\leq 5}$ *AT-EASG* is a close runner-up, outperforming *AT-O2UCT*.

For the largest B_7 games the best-performing method is *AT-O2UCT*, which excels *AT-EASG* (the only remaining competitor) by a clear margin. None of the two exact MILP methods were capable of solving games of this size and the approximate MILP approach (*AT-CBK2018*) solved 16 game instances and failed in solving the remaining 9. Consequently, for the sake of fair comparison, payoff results of *AT-CBK2018* are not presented for B_7 games. Generally, *AT-CBK2018* yields the weakest outcomes across the entire range of game sizes.



(a) Leader's payoff. **(b) Computation time.bib**

Figure 1: (a): The average expected leader's payoff. (b): The average time requirements.

Time scalability is presented in Fig. 1b. While all methods scale exponentially, the running times of non-MILP approaches grow at slower paces. For games from $B_{\geq 6}$ (*AT-EASG*) and from $B_{\geq 7}$ (*AT-O2UCT*), resp. they already excel exact MILP methods. Obviously, the main asset of *AT-C2016* and *AT-BC2015* is convergence to optimal solutions and hence a comparison of their running times with heuristic approaches needs to be considered with care. Nevertheless, it seems reasonable to conclude that beyond certain game complexity the exact methods become infeasible and, in such scenarios, both heuristic approaches present a viable alternative.

The third MILP method is a state-of-the-art algorithm for approximate solving of extensive-form games. Following [4] *AT-CBK2018* was parameterized in a way which assures fast convergence ($\epsilon = 0.3, \sigma = 0.4$), though still for the most complex B_7 games *AT-EASG* and *AT-O2UCT* are faster (Fig. 1b), and at the same time provide better solutions (Fig. 1a). Note that *AT-CBK2018* solved only 16 games from B_7 and times for the remaining instances are capped at the limit of 200h. This situation favors *AT-CBK2018*, as for *AT-O2UCT* and *AT-EASG* the actual times for all games are reported.

Results summary. Evaluation on a set of 25 games shows that non-MILP AT methods (*O2UCT* [7, 8], *EASG* [18]) provide optimal or close-to-optimal leader's payoffs while being visibly faster than exact MILP AT approaches (*BC2015* [2],*C2016* [3]). At the same time, they outperform time-optimized approximate MILP AT method (*CBK2018* [4]) in both payoffs quality and time efficiency.

ACKNOWLEDGMENTS

This work was supported by the National Science Centre, grant number 2017/25/B/ST6/02061.

REFERENCES

[1] Robert J Aumann. 1997. Rationality and bounded rationality. *Games and Economic behavior* 21, 1-2 (1997), 2–14.

[2] Branislav Bosanský and Jiri Cermak. 2015. Sequence-Form Algorithm for Computing Stackelberg Equilibria in Extensive-Form Games. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, Blai Bonet and Sven Koenig (Eds.). AAAI Press, 805–811. http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9610

[3] Jiri Cermak, Branislav Bosansky, Karel Durkota, Viliam Lisy, and Christopher Kiekintveld. 2016. Using Correlated Strategies for Computing Stackelberg Equilibria in Extensive-Form Games. In *30th AAAI*. 439–445.

[4] Jakub Cerny, Branislav Bosansky, and Christopher Kiekintveld. 2018. Incremental Strategy Generation for Stackelberg Equilibria in Extensive-Form Games. In *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, Éva Tardos, Edith Elkind, and Rakesh Vohra (Eds.). ACM, 151–168. https://doi.org/10.1145/3219166.3219219

[5] Daniel Kahneman and Amos Tversky. 2013. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, 99–127.

[6] Jan Karwowski and Jacek Mańdziuk. 2019. A Monte Carlo Tree Search approach to finding efficient patrolling schemes on graphs. *European Journal of Operational Research* 277, 1 (2019), 255 – 268. https://doi.org/10.1016/j.ejor.2019.02.017

[7] Jan Karwowski and Jacek Mańdziuk. 2019. Stackelberg Equilibrium Approximation in General-Sum Extensive-Form Games with Double-Oracle Sampling Method. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '19, Montreal, QC, Canada, May 13-17, 2019*, Edith Elkind, Manuela Veloso, Noa Agmon, and Matthew E. Taylor (Eds.). International Foundation for Autonomous Agents and Multiagent Systems, 2045–2047. http://dl.acm.org/citation.cfm?id=3332005

[8] Jan Karwowski and Jacek Mańdziuk. 2020. Double-oracle sampling method for Stackelberg Equilibrium Approximation in General-sum Extensive-form Games. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20), New York, NY, USA, February, 7–12, 2020*. AAAI Press, 7 pages.

[9] George Leitmann. 1978. On generalized Stackelberg strategies. *Journal of optimization theory and applications* 26, 4 (1978), 637–643.

[10] Jacek Mańdziuk, Jan Karwowski, and Adam Żychowski. 2019. Simulation-based methods in multi-step Stackelberg Security Games in the context of homeland security. (2019). https://sg.mini.pw.edu.pl

[11] Richard D McKelvey and Thomas R Palfrey. 1995. Quantal response equilibria for normal form games. *Games and economic behavior* 10, 1 (1995), 6–38.

[12] James Pita, Manish Jain, Fernando Ordóñez, Milind Tambe, Sarit Kraus, Reuma Magori-Cohen, and Milind Tambe. 2011. Effective Solutions for Real-World Stackelberg Games: When Agents Must Deal with Human Uncertainties. *Security and Game Theory* (2011), 193–212. https://doi.org/10.1017/cbo9780511973031.010

[13] James Pita, Manish Jain, Milind Tambe, Fernando Ordóñez, and Sarit Kraus. 2010. Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artificial Intelligence* 174, 174 (2010), 1142–1171. https://doi.org/10.1016/j.artint.2010.07.002

[14] Ariel Rubinstein. 1998. *Modeling bounded rationality.* MIT press.

[15] Herbert A Simon. 1957. *Models of man: social and rational.* Wiley.

[16] Amos Tversky and Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *science* 185, 4157 (1974), 1124–1131.

[17] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211, 4481 (1981), 453–458.

[18] Adam Żychowski and Jacek Mańdziuk. 2020. A Generic Metaheuristic Approach to Sequential Security Games. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020*. IFAAMAS, 3 pages.