Optimal Risk in Multiagent Blind Tournaments

Extended Abstract

Theodore J. Perkins
Ottawa Hospital Research Institute & University of Ottawa
Ottawa, Ontario, Canada
tperkins@ohri.ca

ABSTRACT

In multiagent blind tournaments, many agents compete at an individual game, unaware of the performance of the other agents. When all agents have completed their games, the agent with the best performance—for example, the highest score, or greatest distance, or fastest time—wins the tournament. In stochastic games, an obvious and time honoured strategy is to maximize expected performance. In tournaments with many agents, however, the top scores may be far above the expected score. As a result, maximizing expected score is not the same as maximizing the chance of winning the tournament. Rather, a "riskier" strategy, which increases the chance of obtaining a top score while possibly sacrificing some expected score, may offer a better chance of winning. In this paper, we study how an agent should optimally adapt its strategy based on the size of the tournament in which it is competing. Our solution involves first approximating the agent's pool of opponents as a collection of known or estimated strategies. Second, score distributions are computed for those strategies, and the distributions are convolved to obtain a distribution for the maximum score of the agent's opponents. Finally, a strategy that maximizes the chance of exceeding the opponents' scores is computed. As a demonstration, we study optimal tournament-size adaptation in online Yahtzee tournaments involving as few as one and as many as ten thousand opponents. We find that strategies change dramatically over that range of tournament sizes, and that in large tournaments, an agent adopting the optimally risky strategy can nearly double its chance of winning.

CCS CONCEPTS

• **Computing methodologies → Stochastic games**; **Dynamic programming for Markov decision processes**; **Agent / discrete models**;

KEYWORDS

multiagent; tournaments; adaptation; optimality; risk; Yahtzee

ACM Reference Format:

Theodore J. Perkins. 2019. Optimal Risk in Multiagent Blind Tournaments. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

1 INTRODUCTION

In many games, agents compete directly against each other and make many tactical and strategic decisions based on their knowledge and observations of their opponents [9, 10, 13]. In other games, however, the agents perform individually, and winning is based on which agent performs the best. Examples include some types of auctions [8], certain game shows, applying for a job or an exclusive fellowship or a grant, and the original motivation for the present work, massively multiagent online tournaments. When many individuals are competing, the bar for success—say, the winning game score in a tournament—is naturally higher. Thus, although each agent is playing individually, an intelligent agent must adapt its strategy based on, at the very least, the size of tournament in which it is competing.

Our original, specific motivation for the present study was the online Yahtzee with Buddies game. For those not familiar with Yahtzee, we describe it in detail below. Briefly, however, it is a turn-based dice game in which agents get points for different combinations of dice. In the traditional, physical version of the game, agents take turns rolling and scoring those roles. They can see each others' roles and scores, and may make strategic decisions based on that information. The Yahtzee with Buddies app has a similar one-on-one mode of play, but it also offers tournament play. In tournament play, each agent plays a game alone, and when all agents have completed their games, the agent with the highest score wins. The app offers tournaments with as few as two other agents and daily tournaments with an unlimited number of agents (typically around ten thousand). In playing such tournaments, it rapidly becomes obvious that the winning strategy for a small tournament differs from the winning strategy in a large tournament. Large tournaments demand risky play, because the score needed to win a large tournament can be exceedingly high. Thus, we became interested in the general phenomenon of risk-taking in blind tournaments, and how optimal play must adapt to the size of the tournament.

Of course, there has been extensive study of risk, risk-reward trade-offs, and related concepts in fields such as game theory, decision theory, economics, and theory of investment [3, 7]. However, we are not aware of prior work specifically on the question of optimal risk as a function of the number of competing agents, and certainly not in the context of our intended application, tournament Yahtzee. Prior work on Yahtzee has mostly focused on the problem of computing strategies that maximize expected score [6, 14, 15] with one work addressing the issue of finding strategies that maximize the probability of exceeding a specific set scores [4]. To our knowledge, only Pawlewicz [11] has seriously study a multi-agent version of the game, with a focus on beating either one or a collection of opponents. They explored heuristic strategies for winning in multiplayer games. In contrast, we focus on computing optimal strategies for winning blind Yahtzee tournaments, and study how the strategy changes as a function of the number of competitors.

In this paper, we propose a general approach to defining optimal agent behaviour in multiagent blind tournaments. To achieve maximum generality, we sought to minimize the assumptions made by our approach. In its simplest form, we require only: (1) that the rules of the game are known; (2) that playing the game results in a discrete-valued score; (3) that the single-agent version of the game can be solved optimally for any terminal objective function that depends on score; and (4) that the number of opposing agents in the tournament is known. Our solution concept also allows the agent to have models of the opposing agents' strategies (as in for example [1, 2, 5]).

2 ALGORITHM

Details of our problem formulation and solution algorithm will appear elsewhere. Briefly, we assume that the single agent version of a game of over interest can be formulated as a Markov decision process [12], and that at the end of the game the agent receives a discrete (say, integer) score. In a *multiagent blind tournament*, one agent plays its own instance of the game, while N other agents do the same. If there is a single agent with a score higher than that of all other agents, then that agent wins the tournament. Our agent may know the policies (i.e. strategies), $\pi_1 \dots \pi_N$, followed by the other agents. Or, our agent may make some default assumption, such as that the other agents each follow a policy that maximizes their own expected score, π_{MES}. In the latter case, our agent must first compute that strategy (using standard dynamic programming techniques), and then set $\pi_1 = \dots = \pi_N = \pi_{MES}$. For each distinct opponent strategy, our agent then uses forward dynamic programming to compute the score distribution for each opponent agent, $Pr(s_i = s|\pi_i)$, where s_i is a random variable describing the final score of opponent agent i, and s ranges over the possible scores. Based on those score distributions, our agent then computes the distribution of the maximum opponent agent score $Pr(\max_{i=1\dots N} s_i = s|\pi_1 \dots \pi_N)$. Finally, our agent uses dynamic programming one more time to compute a policy that maximizes the chance of exceeding the maximum opponent agent score—i.e. the chance of winning the tournament.

3 RESULTS ON YAHTZEE

Yahtzee is a well-known, widely-played dice game, whose rules can be found online (http://www.yahtzee.org.uk/rules.html). In our study of the game, we consider the modification that extra Yahtzees are worth 50 points, not 100, as in the official rules, for consistency with the Yahtzee with Buddies app. Using the app, we manually played tournaments of sizes 15 players, 25 players, 50 players, 5000 to 6000 players, and 10,000 to 12,000 players (the last two being daily tournaments that did not have a fixed size). In these tournaments, we confirmed that winning scores increased dramatically as a function of tournament size, respectively being on average 348.7 ± 41, 352.7 ± 33, 387.9 ± 32, 558.4 ± 23, and 563.4 ± 23. To win the largest tournaments requires multiple Yahtzees, perhaps five or more, which requires a quite different style of play than maximizing expected score.

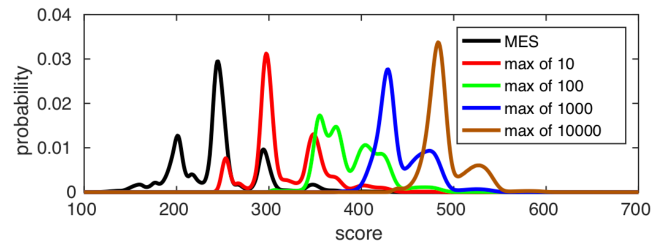

Figure 1: Score distributions for a single player that maximizes expected score (MES), and for the maximum of different numbers of such payers.

We then used our algorithm to compute optimal strategies for an agent to win tournaments against 1, 10, 100, 1000 or 10000 opponent agents, each of which is maximizing its own expected score. This entailed, of course, first computing that score-maximizing policy. Details of how we did this will appear elsewhere. We followed a standard backward dynamic programming approach [12], using a state space similar to that used in previous analyses of Yahtzee [4, 6, 14, 15], but including the exact game score (which is not necessary when one is maximizing expected score). We solved the dynamic program using bespoke parallel code, running up to 50 processes in parallel, in approximately one week on a Compute Canada cluster at the Centre for Advanced Computing at Queens University. A forward dynamic program on the same state space was then used to compute the distribution of final scores for an agent following that policy (black curve in Figure 1). From that we computed the distributions for the maximum score of 10, 100, 1000 or 10000 such players (other curves in Figure 1)—the "score to beat" for our intended tournament-optimal player. These score distributions are interestingly complex, with many distinct modes corresponding to whether or how many Yahtzees are scored, along with certain other high-scoring boxes, such as large straight. Finally, we computed optimal agent strategies for winning tournaments against these different numbers of opponent MES agents. Details of that policy will appear elsewhere. We found that the tournament-optimal strategy, as opposed to playing the MES strategy, could increase the chance of winning the tournament by 2.67% (against 1 MES player), 9.44% (against 10), 30.5% (against 100), 49.9% (against 1000) and 69.7% (against 10000). The tournament-optimal strategy does this by adopting increasingly riskier strategies—essentially trying to score as many Yahtzees as possible—with expected score decreasing respectively as 241.6 ± 43, 234.3 ± 52, 222.1 ± 56, 209.1 ± 60, 198.3 ± 61.

4 CONCLUSION

Optimal play in multiagent blind tournaments can be solved by a series of dynamic programs. At least in the case of Yahtzee, optimal play in tournaments of increasing size—in the sense of maximizing the chance of winning the tournament—is associated with riskier play. Decreased expected score, but with higher variance, gives a better chance of obtaining a high, tournament-winning score.

REFERENCES

[1] Darse Billings, Aaron Davidson, Jonathan Schaeffer, and Duane Szafron. 2002. The challenge of poker. *Artificial Intelligence* 134, 1-2 (2002), 201–240.

[2] David Carmel and Shaul Markovitch. 1996. Learning models of intelligent agents. In *AAAI/IAAI, Vol. 1*. 62–67.

[3] Louis Anthony Cox, Jr. 2009. Game theory and risk analysis. *Risk Analysis: An International Journal* 29, 8 (2009), 1062–1068.

[4] CJF Cremers. 2002. How best to beat high scores in Yahtzee: A caching structure for evaluating large recurrent functions. *Master's thesis, Fac. of Math. and CS, Technische Universiteit Eindhoven, The Netherlands* (2002).

[5] Sam Ganzfried and Tuomas Sandholm. 2011. Game theory-based opponent modeling in large imperfect-information games. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 533–540.

[6] James Glenn. 2006. An optimal strategy for Yahtzee. *Loyola College in Maryland, Tech. Rep. CS-TR-0002* (2006).

[7] Christian Gollier. 2004. *The economics of risk and time*. MIT press.

[8] Vijay Krishna. 2009. *Auction theory*. Academic press.

[9] Roger B Myerson. 2013. *Game theory*. Harvard university press.

[10] Martin J Osborne and Ariel Rubinstein. 1994. *A course in game theory*. MIT press.

[11] Jakub Pawlewicz. 2010. Nearly optimal computer play in multi-player Yahtzee. In *International Conference on Computers and Games*. Springer, 250–262.

[12] Martin L Puterman. 2014. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

[13] Lloyd S Shapley. 1953. Stochastic games. *Proceedings of the national academy of sciences* 39, 10 (1953), 1095–1100.

[14] Tom Verhoeff. 2000. Optimal solitaire Yahtzee strategies.

[15] Phil Woodward. 2003. Yahtzee®: The Solution. *Chance* 16, 1 (2003), 18–22.