Research Project Proposal

Mike Ledger <u5582972@anu.edu.au>

April 17, 2023

1 Introduction

1.1 Background; Tournaments in sports and eSports

Sport and eSport tournaments have massive prize pools and ostensibly aim to determine the best players or teams within their respective game. According to statistics website Statistica, the total sum of prize money across the largest 10 sports was approximately $6bn USD. According to Esports Earnings, in 2022 $229,675,986.73 USD was spent in eSports prize pools. While are not figures for the total expenditure in prize money across all sports and eSports tournaments, they can be used as lower bounds; which are evidently very large.

It is thus important that tournament events should use fair methodologies that do maximise the fairness of the event, how reliable the result of the event is in ranking players against each other, and for tournaments to be robust against possible manipulation by players. "Manipulation" can also happen unintentionally; players may simply forfeit games in multi-stage or double-elimination elimination tournaments, due to whatever valid circumstance. This can lead to their result at the end of one tournament stage being lower than it would have otherwise been, and thus result in an unfair position going into the next stage of that tournament. [1]

For the purpose of this proposal the term, a "game" shall mean a single sports match result or eSport game result, and a "player" shall mean a participant (contestant, player, participant, etc.) in a "game".

1.2 Quantifying the aims of a tournament

Creating metrics that quantify these attributes has been done with various approaches. Fairness has many possible definitions; such as that a tournament should not be onerous for stronger players, it should not ultimately disadvantage a player for winning any game, the rules are applied consistently, and importantly: that the stronger players make it further than the worse players [1].

The reliability of a tournament is also a key question in its efficacy; did the truly strongest player win? Did the next strongest players come in the next rankings, in order of skill? Would the outcome of the tournament be significantly different under a different structure? With what precision are

[1] In such cases, it seems contrary to fairness to either drop the player completely from the tournament, due to a circumstance out of their control leading to their forfeited game(s), or alternatively to award them wins despite having forfeited, or alternatively to accept the forfeits as losses and thus artificially lower their rank. In general, these trade offs are hard to manage.

players ranked? [2] attempts to answer the latter question and proposes a novel tournament structure with excellent **ranking precision** across all the final rankings, and [3] finds that Swiss-style tournaments are more reliable than knockout-style tournaments in terms of reliably producing rankings that reflect the true skill of each participant.

A key term to be aware of in terms of tournament structure is the **seeding** of a tournament. This refers to the initial ranking of each player at the outset of the tournament; the strongest player would receive seed 1, and the weakest seed n. The seeding of a tournament is strongly determinative of its outcome, especially in single-elimination (also known as knockout) tournaments, and to a lesser extent in double-elimination (variously named "draw and process", double knockout, etc.) tournaments [4] [5] [6].

1.3 Exploring existing tournament structures

As hinted at above, many different tournament structures are possible, and each has different characteristics in terms of the number of games required.

These include: how many rounds those games must be played in (i.e., how parallelisable the games are); the ranking precision [2]; fairness; competitive development or spectator interest; among many others.

The purpose of this proposal then is to explore the efficacy of novel tournament structures in terms of those metrics outlined above. My personal stake in this field is that I have created and run a tournament website [7] focused on Arena FPS games such as Quake Champions. Over 3000 games have been played through this website over ~180 separate tournaments since its inception in 2020. But the question of what tournament structures should be available is not easy to answer, especially in light of imperfect seeding. Time constraints are also a key consideration; as this website facilitates mostly amateur-level competition, players are both unwilling and unable to devote more than 2 or 3 hours to a single tournament - they do not have the time.

For example, balanced single-elimination tournaments give good ranking precision for 1st place only, and require only $\log_2(n)$ rounds to be played [3]. Which player becomes the 2nd place finisher of a single-elimination tournament is largely determined by seeding; and similarly (but to an even greater extent) the same is true for 3rd place, and so on. Indeed, only the 1st and 2nd place results have unique rankings in single-elimination. Viewing the winning position of a single elimination as the root of a tree, at each level above it, half of the contestants are eliminated; so each eliminated player has a final ranking isomorphic to the round they were eliminated in. See Figure 1.

Alternatively, double-elimination tournaments give much better [5] ranking precision for the 1st through to 4th places, and control for both "upset" losses and for imperfect seeding better by granting players two losses in total (instead of just one) before they are eliminated from the tournament. However, double elimination tournaments require $2\log(n)+1$ rounds. As we see in Figure **??**, in terms of number of rounds required, this is not an improvement over round-robin tournaments for $n \leq 6$, while also having less ranking precision by simply having fewer unique rankings at the end of a tournament.

Format	Uniq. rankings	Rounds; $R(N) = \cdots$	$R(2)$	$R(3)$	$R(4)$	$R(5)$	$R(6)$	$R(7)$	$R(8)$
R.R.	n	$n-1$	2	3	4	5	6	7	8
S.E.	$\lceil\log(n)\rceil$	$\lceil\log(n)\rceil$	2	2	3	3	3	3	4
D.E.	$\lceil\log(n)\rceil * 2$	$\lceil\log(n)\rceil * 2 + 1$	4	4	6	6	6	6	8

However, the asymptotic growth of round robin stages quickly make them too large to run for

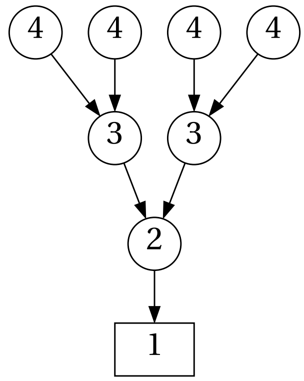


Fig. 1. Shared rankings of players at each level of a single-elimination tournament with $n = 8$, formulated as the level of the tree starting from the winning rank node. Each oval represents a match, and each node is labelled with the rank (possibly shared) of the player at that match if they make it that far.

increasing n, while they also do not incentivise player or spectator interest in each individual game; the consequence for both winning and losing a game in a round-robin is lower than for knockout-type tournaments.

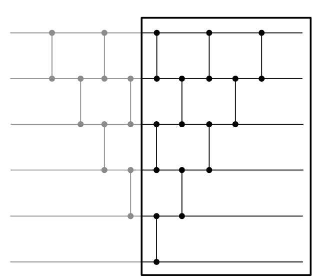
Swiss-style tournaments offer an interesting middle-ground. Each round pairs players with roughly the same score, though two players cannot meet more than once [3]. The choice for this pairing algorithm is usually left to a computer implementation, as it must ensure that the no-rematch policy is met. Indeed, [3] shows that Swiss-style tournaments produce more reliable results than knockout tournaments using real-world player rankings.

2 Research Problem

These are just examples. There are many more possible tournament structures, and this proposal is in particular interested in the characteristics of compositions of tournament structures, as well as exploring the characteristics of using sorting algorithms as a tournament structure. The analogy here is that each game in a tournament is a comparison in a sorting algorithm.

In particular, many existing popular tournament formats can be modelled as sorting networks; or vice-versa, and comparing existing tournament formats to known sorting algorithms can reveal some interesting comparison.

We can show that a single-elimination tournament has the structure of a the latter half of an insertion sort or bubble sort (if allowing for parallel comparisons in each "round" of either sorting algorithm) [8]:



Then what are the characteristics of using parallel insertion/bubble sort as a tournament algorithm? What of bitonic sort? Or merge sort? Indeed, we have already seen a tournament format that involves a complex pairing algorithm (Swiss-style). While many comparison-based sorting algorithms are within $o(n \log n)$ time complexity, as seen in 2, we can strategically limit the scope of the algorithm to achieve an acceptable time complexity, for instance by limiting the number of depth of a quicksort.

The aim of this project would be to answer these questions to analyse various tournament structures. This would be accomplished by the design and implementation of a tournament structure domain-specific language to enable the definition, and simulation of, various tournament structures.

The purpose of this research would be to create a minimal computer language for the definition and analysis of tournament structures.

This language, would enable definition, simulation, and analysis of novel tournament structures such as those defined by sorting algorithms, compositions of existing tournament structures (such as multi-stage tournaments; a common format being group stages of g players per group, where the top k players then advance into a final elimination stage).

The language would be accompanied by a tool that shall be able to take as input the size n of a tournament (in players), and their skill distribution by Elo [9] ratings, and use those ratings to perform simulations and determine the likely outcomes of running a tournament with that structure and rating distribution.

This can be used to see how effectively the tournament preserved the apriori skill of a player through the tournament.

The following formats are proposed for analysis at the outset, and thus must be supported for definition within the language:

1. Single-elimination (balanced and equal gap [4] tournaments)
2. Double-elimination
3. Round robin
4. Shell-sort
5. Quick-sort
6. Insertion-sort
7. Reaper [2]

With two additional special compositional formats:

1. Groups of G as a composition of any above format
2. Sequences of G as a composition of any above format

The advance in the literature that this project will provide is to create a simple compositional notation for defining tournament structures, and the result of the simulation and analysis which shall be able to provide measurement of the quantifiable tournament characteristics (see the introduction) across a variety of possible structures.

A significant contribution of the project will be to compare sorting algorithms (fully or partially applied) to popular tournament structures in terms of their efficacy as tournaments. This may have application beyond tournament design; in machine learning the problem of producing a ranking using noisy comparisons is an open field of study [3], [10].

3 Methodology

The method for completing the research project is defined here.

3.1 Literature Review

A thorough literature review encompassing:

- Near, noisy, stochastic, probabilistic, etc. sorting algorithms
- Sorting networks
- Tournament structures

Shall be performed in order to find the current state of the relevant fields. This review will inform the focus of the project so as to not duplicate work that has been done before. The review will not be a major output of the project as there exist already recent articles that thoroughly review the existing relevant literature already [1], [2], [4], [5], [11].

3.2 Domain-specific language (DSL) design and implementation

Review shall also be done on literature concerning DSL design to inform the design of the language. A grammar shall be constructed and iterated upon for the language's interpreting tool to use.

3.3 Tool simulation implementation

The tool shall be augmented to be able to perform simulation of the input tournament structure, given a list of Elo ratings. Simulations shall be ran that, on aggregate, determine how well the input tournament structure performed according to the quantifiable metrics (for instance, reliability can be encoded as number of inversions in the final rankings [3]).

3.4 Tournament structure definition

The structures outlined in the Research Problem shall be defined, and simulations performed using them. Additionally, sorting algorithms shall be implemented under the DSL.

3.5 Overall analysis

Analysis shall be performed to show which tournament structures performed the best according to the defined metrics. A report shall be written that outlines these findings, the DSL defined, and the definitions of tournament structures under the DSL.

4 Evaluation Criteria

There are a variety of evaluation criteria for this project;

1. The success of the literature review in reducing any duplicate work the project may perform.

2. The success of the DSL in defining existing and novel tournament structures. Ideally, a notation is found that is not simply a Turing-complete programming language, but reveals a deeper structure of underlying sorting network-like structures (such as most tournament structures, with the exception of Swiss-style).

   Alternatively, a finding is made that reveals the impossibility of creating such a language that is able to define popular tournament structures *without* a Turing-complete programming language.

3. The scope of the final analysis; what formats are analysed? How many formats are analysed? Are novel tournament structures identified? Are the characteristics of these novel tournament structures shown to be better or worse than existing formats?

5 References

[1] P. C. Placek, "The impossibility of a perfect tournament," *Entertainment computing*, vol. 45, p. 100540, Mar. 2023, doi: 10.1016/j.entcom.2022.100540.

[2] N. P. H. Bao, S. Xiong, and H. Iida, "Reaper tournament system," *Intelligent technologies for interactive entertainment*, pp. 16–33, 2018, doi: 10.1007/978-3-319-73062-2_2.

[3] B. R. Sziklai, P. Biró, and L. Csató, "The efficacy of tournament designs," *Computers & operations research*, vol. 144, p. 105821, Aug. 2022, doi: 10.1016/j.cor.2022.105821.

[4] A. Karpov, "A new knockout tournament seeding method and its axiomatic justification," *Operations research letters*, vol. 44, no. 6, pp. 706–711, Nov. 2016, doi: 10.1016/j.orl.2016.09.003.

[5] I. Stanton and V. V. Williams, "The structure, efficacy, and manipulation of double-elimination tournaments," *Journal of quantitative analysis in sports*, vol. 0, no. 0, pp. 1–17, Jan. 2013, doi: 10.1515/jqas-2012-0055.

[6] L. Csató, "Quantifying the unfairness of the 2018 fifa world cup qualification," *International journal of sports science & coaching*, vol. 18, no. 1, pp. 183–196, Apr. 2022, doi: 10.1177/1747954121107

[7] M. Ledger, "Kuachi cups." 2020. Available: `https://kuachi.gg`

[8] D. Knuth, *The art of computer programming: Sorting and searching, volume 3*. ADDISON WESLEY PUB CO INC, 1998. Available: `https://www.ebook.de/de/product/3236573/donald_knuth_the_art_of_computer_programming_sorting_and_searching_volume_3.html`

[9] A. E. Elo, "The proposed uscf rating system - its development, theory, and applications," 1967.

[10] W. Ren, J. Liu, and N. B. Shroff, "On sample complexity upper and lower bounds for exact ranking from noisy comparisons." arXiv, 2019. doi: 10.48550/ARXIV.1909.03194.

[11] D. Van Bulck and D. Goossens, "Handling fairness issues in time-relaxed tournaments with availability constraints," *Computers & operations research*, vol. 115, p. 104856, Mar. 2020, doi: 10.1016/j.cor.2019.104856.