




Concerning the Rules of Tournaments

his article is to some extent a reply to Γιώργος Ζαχαρόπουλος's article¹ in which he points out the conflicting requirements that tournament rules have to address. Ζαχαρόπουλος examines tournaments under three aspects: safety, scoring and spectacle.

Safety certainly is a crucial aspect as we want to simulate a process of mutual harming and killing while avoiding actual injuries. Regarding the set of rules, the question arises if possibly restricting or unrealistic equipment should be prescribed and if dangerous attacks such as targeting the hands or other insufficiently protected areas should be banned. Safety is a matter of discretion and while I think that a compromise between safety and closeness to reality has to be found in a tournament setting, I cannot define exactly what this compromise would be and thus will not address the question of safety any further in the scope of this article.

I also do not want to consider the tournament as a spectacle. In my opinion, a tournament is primarily for the contestants and those who want to learn from the contestants. While I do not object to an audience enjoying a tournament, I do not think that any adaptations should be made in order to make a tournament more entertaining if these changes make the fights less realistic than they could be.

1 Tournaments as a Training Tool

Instead of considering safety and spectacles, I will focus on fair scoring and on another aspect, tournaments as a training and evaluation tool. Despite this article being written with training in mind, it will probably apply to most contexts in which a tournament may take place.

Comments on Ζαχαρόπουλος's article suggest that a variety of different sets of rules for different tournaments impedes the exploiting of flaws or undesired discrepancies between a

¹<http://www.hroarr.com/on-tournament-rules/>

real fight and a tournament match. The argument is that a good fencer, i. e. a fencer who is likely to survive the combat situation which the tournament should simulate, is also more likely to win a tournament match when the contestants do not have the opportunity to find or exploit the flaws. However, a flawed set of rules does not necessarily need to be analysed and willingly exploited in order to produce flawed results. If the results of a tournament should reflect the results of real fights, then the rules need to reflect the rules of real fights.

A real fight in the context of historical martial arts may occur in different situations, such as a trial by combat, an assault or on a battlefield. Depending on the particular sort of encounter that the tournament matches simulate, different rules may apply: The match may be held on tidy even ground or between obstacles, with a time limit or without, with the option of escaping the fight or not, to the death or first blood etc.

At Tremonia Fechten, we use a variant of Hammaborg's HALAG set of rules², see the last section of this article. While the analysis of tournament rules in the scope of this article is a general one, our modified HALAG³ rules are often referred to as they are the result of the analysis.

2 Referees

The HALAG rules suggest to not rely on referees, but instead let the contestants themselves count and score the hits they suffer. There are several advantages to this approach. The fencers are encouraged to pay attention to their opponents' hits instead of focussing on only their own attacks, which is particularly valuable when tournaments are considered as a training tool. Furthermore, the fencers have a tactile perception of the hits they take that a referee lacks. Also, if no referees are required, the availability of referees does not limit the number of simultaneous matches, which is useful in tournaments with many contestants.

On the other hand, leaving the scoring to the fencers themselves may result in an imbalanced or biased perception of the hits or even in "undead" contestants. From our experience, however, hits that would be counted by a neutral referee but are ignored by the receiving fencer are an exception. This might be due to the social pressure among members of the same club, so we do not know yet if contestants that do not see each other frequently would also comply and fairly score the hits they suffer.

3 Single Match

A contestant can survive a match unharmed, get harmed or get killed. These are the possible outcomes that are considered in Tremonia Fechten's exegesis of the HALAG rules. A finer graduation between unscathed and dead is also possible, but it should be kept in mind that a finer graduation of the fencer's state of health requires a more nuanced grading of the effect of a hit. It should also be considered that it is quite difficult to judge

²<http://www.anno1838.de/publikation>

³Hammaborg Laurentiusgildet Annual Gathering

the cumulative effects of several non-lethal hits and, as well, perform adequate in-fight reaction to the blows a fencer would take during a fight. Therefore, some simplification is necessary, which resulted in the dead/injured/unharmd scheme according to the simple final/harming hit pattern of the original HALAG rules.

It makes sense that a fencer's primary aim is to survive the fight, thus survival should be rated highest in a tournament setting. Depending on the exact set of rules, survival might be ensured by different means, e. g. killing the opponent or escaping the fight. The original HALAG rules are very flexible in this point and allow considerable modification in order to mimic the combat situation that the tournament should reflect.

4 Round-Robin or Elimination?

We see the tournament as a means of evaluation, testing and training of the competing fencers. Although tournaments are often entertaining as well, entertainment is not the primary purpose. Let us compare the two tournament archetypes, single-elimination and round-robin, in the context of their respective training and evaluation values. Clearly, a single-elimination tournament in which the best fencers get the most fights does not serve the training objective very well. Additionally, single-elimination is meant to find only the best fencer while the other positions remain uncertain. In contrast, a round-robin tournament not only allows each of the n contestants the same number $n - 1$ of fights, it also allows the ranking of all the participants, and not only the best one. In a single-elimination tournament, even the process of finding only the best fencer is questionable. Single-elimination relies on two premises:

1. The better fencer wins the fight.
2. Being better than the other fencer is transitive, i. e. if Alice beats Bob, and Bob defeats Charlie, then Alice beats Charlie.

Looking at the data from round-robin tournaments, we see that both these premises are wrong in their absoluteness. Instead, we most often see that even the fencer who survived the most matches was killed in at least one match. This means that the fencer with the highest probability of survival < 1 can bow out during any round of a single-elimination tournament.

Practically, the number of concurrent fights is limited by the space available and, if referees are considered necessary, the number of referees. Also, there is only a certain amount of time for the tournament. Thus, depending on the number of contestants n and the circumstances, the number of matches might be a factor. If we look at the number of matches r in a round-robin tournament, we see that the first contestant Alice will have to fight against all the other $n - 1$ contestants. After that, second contestant Bob will fight all the contestants except himself and Alice, whom he already fought before, i. e. $n - 2$ matches. Charlie then will fight against the $n - 3$ contestants he has not yet fought etc. The last contestant will have already fought each of the other contestants and thus has no more matches to compete in, i. e. 0. Hence, the total number of matches r in a round-robin tournament is the sum of the whole numbers from 0 to $n - 1$, i. e.

$r = 0 + 1 + 2 + \dots + (n - 2) + (n - 1)$. In a single-elimination tournament, the total number of matches s is $n - 1$, as each match produces one loser. In order to determine the best fighter from n contestants, there have to be $n - 1$ losers. If we compare the numbers of matches

$$r = \sum_{i=0}^{n-1} i = \frac{n(n-1)}{2} \quad (1)$$

$$s = n - 1 \quad (2)$$

we see that a round-robin tournament results in $r/s = n/2$ as many fights as a single-elimination tournament, see Fig. 1. Also, during a single-elimination tournament with n participants, the individual fencer has to engage in only $\lceil \log_2 n \rceil$ fights or less, while in a round-robin tournament, there are $n - 1$ fights for each contestant, see Fig. 2.

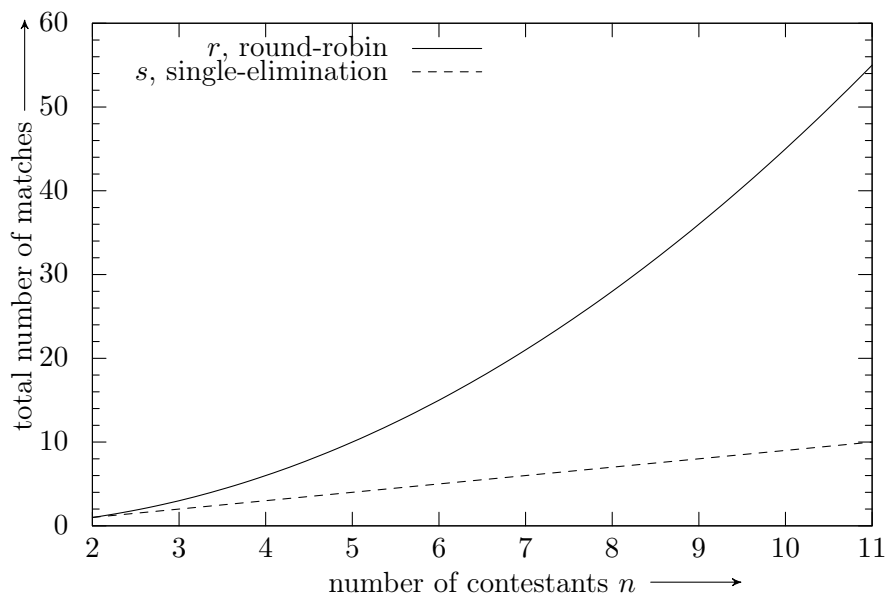


Figure 1: Total number of matches for n contestants in a round-robin vs. a single-elimination tournament

The larger number of total and individual fights in a round-robin tournament requires a match to be considerably shorter than a match in a single-elimination tournament. This comes in handy as, as described above, a match must be short in order to reflect a real fight. For tournaments with very many participants, however, a round-robin system might still not be applicable depending on the circumstances. Several tournament types offer a compromise between the number of matches and the precision of the result.

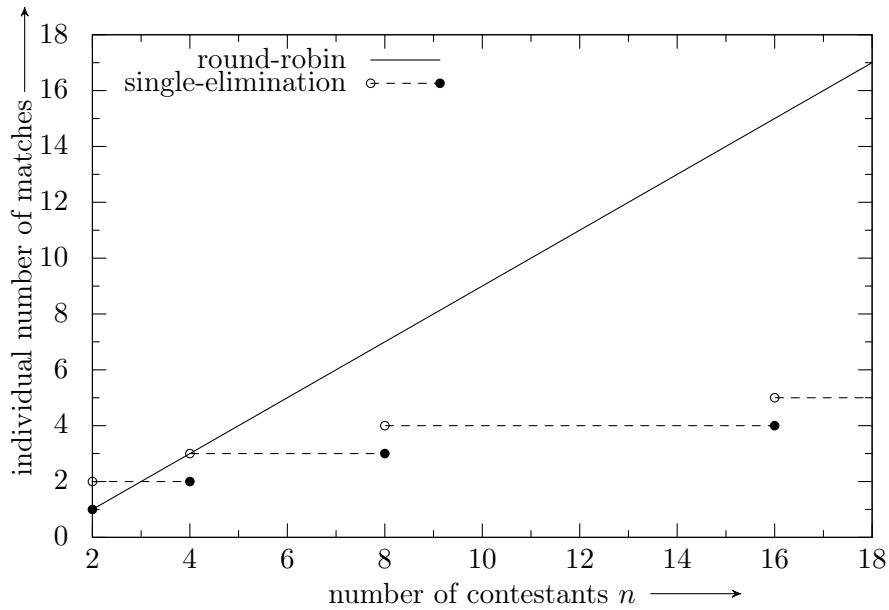


Figure 2: A contestant's maximum number of matches in a round-robin vs. a single-elimination tournament with n contestants

5 Scoring a Match and Evaluating the Result

A match ends with contesting fencers counted as unharmed, dead or, depending on the rules, some intermediate state. In a single-elimination tournament, the loser naturally bows out while the winner advances etc. until the best fencer is found. However, for the aforementioned reasons, elimination modes should be avoided unless the number of contestants is too large for a round-robin tournament. While a single-elimination tournament is straightforward to evaluate, a just and comprehensive evaluation process for round-robin tournaments has to be determined.

We often see that a contestant is awarded a certain amount of points after each fight depending on either the fencer's or the opponent's state of health. For instance, let us assume that an unharmed fencer is awarded 2 points, a harmed but surviving fencer is awarded 1 point and a dead fencer gets 0 points; after the tournament, the points are summed up for each fencer, and the contestant with the most points makes the first place, the contestant with the second-most points makes the second place etc. At first glance, it seems natural to award the best result, i. e. surviving a fight unscathed, the most points and the worst result, getting killed, the least points and then simply add up the points and compare. However, this scoring method has some strange implications which make it, to put it mildly, debatable. Let us regard the above example scoring. It implies that

	A	B	C	D	Σ_o
A	—	(1, 0, 0)	(1, 0, 0)	(0, 1, 0)	(2, 1, 0)
B	(0, 0, 1)	—	(1, 0, 0)	(1, 0, 0)	(2, 0, 1)
C	(0, 1, 0)	(0, 1, 0)	—	(1, 0, 0)	(1, 2, 0)
D	(1, 0, 0)	(0, 1, 0)	(0, 0, 1)	—	(1, 1, 1)
Σ_c	(1, 1, 1)	(1, 2, 0)	(2, 0, 1)	(2, 1, 0)	

Table 1: Evaluation of an example tournament

surviving one fight unharmed and dying in another fight is as good as surviving both fights harmed. Odd implications like this can be found for any tuple of points awarded.

Instead of awarding points which inherently weigh the considered conditions, the conditions themselves should be noted down. Let us look at a very small example tournament of four fencers, Alice (A), Bob (B), Charlie (C) and Dana (D), with three possible states of the fencers after each match: dead, injured, unharmed. The occurrences of these states are written down for each contestant and summed up as a status tuple. For instance, let us write *dead* as (1, 0, 0), *injured* as (0, 1, 0) and *unharmed* as (0, 0, 1). We then write these tuples below the respective contestants and sum the columns, see Tab. 1. Dana (D) got killed by Bob (B) and Charlie (C) and injured by Alice (A). She never survived unharmed. Thus, Dana’s combined status tuple $\Sigma_c(D)$ is (2, 1, 0). How do we define the best fencer then? While the elimination criterion – the one who survives the tournament is the best contestant – is usually not applicable in a round-robin tournament, it would be appropriate that the one with the highest survival probability would be the best fencer. The one who survives most often is the one who gets killed the least, so sorting the tuples in ascending order by their first element – how often one died – will result in the tuples that represent the highest survivability showing up first. In our example, Alice as well as Bob were killed only once each and Charlie and Dana each died twice. We find that Alice did better than Bob, because Alice got a non-lethal injury only once, while Bob suffered a non-lethal injury twice. Analogously, Charlie did better than Dana, so sorting the status tuples Σ_c similar to the way in which words are sorted in a dictionary yields the ranking:

1. Alice: (1, 1, 1)
2. Bob: (1, 2, 0)
3. Charlie: (2, 0, 1)
4. Dana: (2, 1, 0)

The same sorting method can be used if more than one level of injury between *dead* and *unharmed* is considered. A peculiarity of this ranking method is that it can be applied even if the contestants engaged in different numbers of matches, e. g. when a fencer could not continue because of a real injury or if a year of monthly tournaments has to be evaluated.

Instead of sorting by the total numbers of the respective statuses, the results should then be sorted by the *normalised* tuple elements; e. g. Dana contested three matches, so we divide her contestant's status tuple $\Sigma_c(D) = (2, 1, 0)$ by 3 in order to get her status probabilities: $\Sigma_{c,n}(D) = (\frac{2}{3}, \frac{1}{3}, 0)$. We see that Dana's probability of dying is $2/3$, her probability of surviving harmed is $1/3$ and her probability of surviving unharmed is 0. There lies another advantage of this scoring method: Instead of a ranking based on mostly meaningless points derived from an arbitrary weighing process, we have a ranking based on concrete performance data.

According to the original HALAG rules, a fight is ended by either dying of at least one of the contestants or by escaping the wide measure. If a match can end without a contestant being killed, survival probabilities might be too high to establish a ranking based solely on the contestants' status tuples Σ_c . On these occasions, it makes sense to not only take into account the contestants' status tuples Σ_c , but also the respective opponents' status tuples Σ_o . As it is usually better to kill an opponent in a combat situation than leaving him alive, the opponents' status tuples Σ_o should be sorted in *descending order*. Again, the status tuples should be normalised when the contestants fought different numbers of fights or when probabilities instead of sums are to be evaluated. From these tuples Σ_o , additional data such as a fencer's deadliness can be deduced.

6 Summary

Referees are not always necessary for a tournament. Depending on the social pressure on the contestants and their attitudes, referees might be omitted and replaced by the contestants themselves who count and score the hits they suffer. If a good fencer is defined as a fencer who would likely survive a real combat situation, then the matches must be as close to reality as reasonably possible. This means that a match takes only as long as it would take until a fencer would be considered dead or injured, resp., depending on the simulated encounter. Therefore, realistic matches would usually be short, which in turn allows to conduct numerous matches. Numerous matches are necessary in a round-robin tournament which, other than a single-elimination tournament, ranks every contestant and not only the best one. In addition to a precise ranking, a properly evaluated round-robin tournament gives valuable performance feedback for the contestants.

Acknowledgement

I would like to thank the head of Anno 1838, Marcus Hampel, who is the maintainer and one of the creators of the original HALAG rules, for patiently discussing the intentions and implications of the set of rules with me. Thanks are due to Patrick Gerhold and Julian Uszkoreit who introduced the HALAG rules as the basis of our club tournaments, and thanks to Daniel Heumann and Jan Hoffmann for their input on the statistics-based ranking.

Appendix: Tremonia Fechten's Tournament Rules

1. The match starts after saluting as soon as the fencers enter wide measure.
2. The match ends as soon as at least one fencer counts as *dead* or *unable to fight*, which he or she indicates by distinctly stepping back.
3. A fencer counts as *dead* or *unable to fight* when he or she suffers at least one lethal or two non-lethal harming hits. Harming hits are hits that would not immediately terminate the fight but would inflict an injury that requires medical attendance if executed with a sharp sword and without protective gear. Lethal hits are hits that would immediately kill an opponent if executed with a sharp sword and without protective gear. These are, in particular,
 - (a) strikes or thrusts to the head,
 - (b) strikes or thrusts to the neck,
 - (c) thrusts to the chest.
4. Immediately after getting hit, a *dying* fencer may finish his or her attempt to hit the opponent if the attempt has already started.
5. Each fencer counts and scores the hits he or she suffers and notes down the result. Referees are not required, but may be consulted in ambiguous situations.
6. The fencers note down whether they have been killed, harmed or remained unharmed. The survival probability determines the ranking. If two fencers or more share the same survival probability, they are ranked in order of their probability of staying unharmed.