Project Proposal

Project Title: Using Computer Programming and Reseeding to Create Fair Horseshoe Tournaments

Author: Ian Grzembski

Date: October 26th, 2020

Project Definition:

The overall aim of this project is to create a dynamic tournament matching system that responds to people's performances in it. Horseshoe tournaments in particular will be analyzed, and ringer percentages (average skill) and day-to-day differences from their average will be considered. I expect that considering these two factors might affect the smoothness of the tournament, but they will make it more fair.

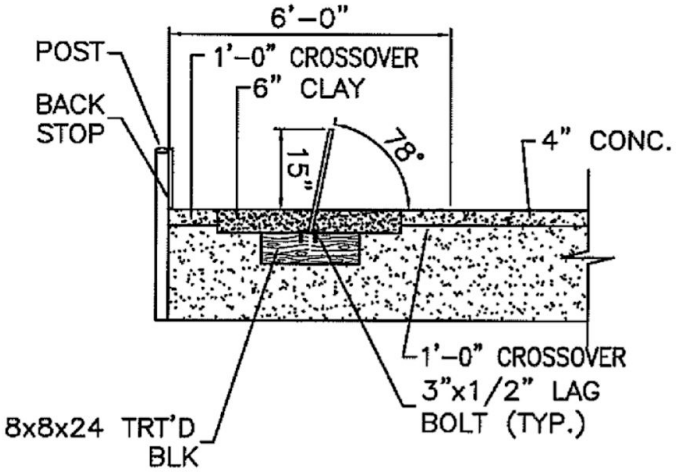
Phrase 1: How can a tournament be arranged such that it is fair for all participants playing, considering their performance during the tournament and previous skill?

Phrase 2: The goal of this project was to develop an algorithm that manipulates the structure of tournaments to make matchups fair, equitable, and exciting for the audience.

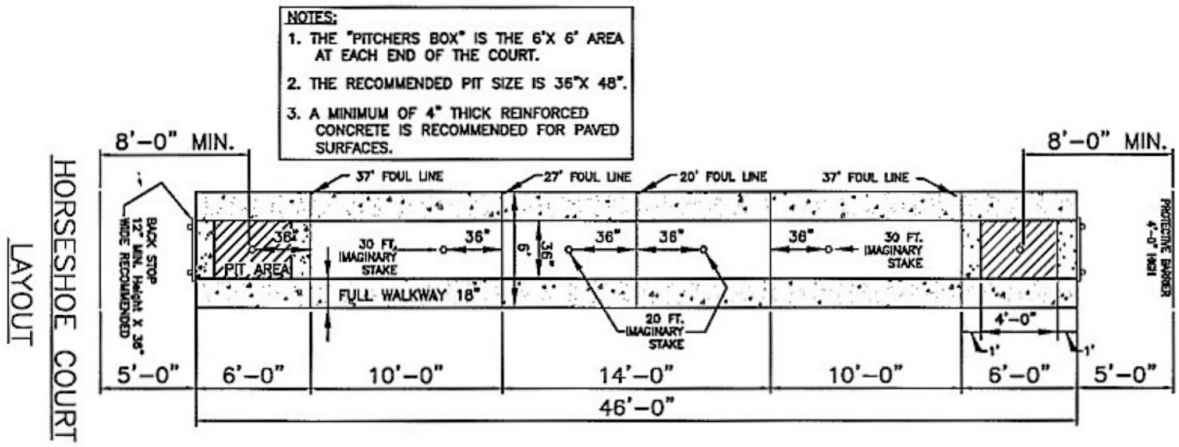
Background:

The Basics of Horseshoes

Horseshoes is a skill based game that has teams of two. Two teams normally play on what is called a horseshoe pit (which is sometimes called a court) which is filled with sand, dirt, or clay. The diagrams below show an example set of horseshoe pits with labels (NHPA, 2019, 49-50). A horseshoe court has two sides or pits, on each there is a metal pole, the poles are 40 feet apart. Men must throw horseshoes from behind the 37' foul line, while women, children, and the elderly (with approval) can throw from behind 27' (NHPA, 2019, p. 8-9).



A side view of a horseshoe pit with distance and angle labels. (NHPA, 2019, 49)



A bird's eye view of a horseshoe pit, with distance labels. (NHPA, 2019, 50)

Both teams try to score the most points by throwing their horseshoe onto the pole without hitting any edge of the pit to score a ringer, which is worth 3 points. If the horseshoe does not come to rest around the pole, but is still 6 inches (or one horseshoe's width) away from the pole, it counts as one point (NHPA, 2019, p. 11).

There can be multiple objectives to win a game of horseshoes. Some games end when one team reaches a certain number of points (21, 40, etc.), some games end when a certain even number of horseshoes are thrown (40, 50, etc.), and others end when either of these conditions are met, whichever one came first (NHPA, 2019, p. 17). For example, in the Central Connecticut Horseshoe League (CCHC), games consist of 15 frames. In each frame, both members of each team throw 2 shoes. The highest score at the end of these frames wins the game (Sluys).

Ringer Percent

Ringer percent is an average skill level currently used to rank people in more professional tournaments. It is simply the number of ringers a person throws out of how many shoes they throw (NHPA, 2019, 29). This does not count any shoes that do not become ringers, even if they were very close to that point.

The Current Problem with Horseshoe Tournaments:

Currently, partners in horseshoe tournaments are chosen by random draw. If every player was the same skill level, this would be fair, however when there is a big gap in skill, this becomes problematic.

Horseshoe tournaments can be unfair. When two very skilled players become partners and easily beat the competition, it is not very enjoyable for everyone who has lost an opportunity of winning to a random draw. This might cause some players to stop trying, whether that be the "top players" sympathizing for a lower skilled friend that was kind to them that they have to play against (Rabin, 1993, p 1281), or a lower skilled player not trying because they know they will lose.Even if two "top players" do not become partners, they may end up having a partner who has won previously, making the tournament not as interesting for the fans.

There is also no seeding system in place, making it more difficult for fans to guess who might win without prior knowledge of who has more skill. A seeding system is a bracket in a tournament where each team is assigned a numerical seed or skill level out of how many teams there are. For example, the seed 1 is the team with the most skill, the seed 2 is the team with

the second most skill, etc. (Khatibi, King, Jacobson, 2014, p. 1). Even if a seeding system were to be implemented, there would still be issues. The standard method of seeding a tournament (in a 16 team tournament, seed 1 plays seed 16, seed 2 plays seed 15 on the other side of the bracket, seed 3 plays seed 14, seed 4 plays seed 13 on seed 1's side of the bracket, etc.) is flawed in the second round, as lower seeds can have a higher chance of winning than a higher seed, which is not supposed to happen (Schwenk, 2000, 6-8). The standard method, in effect, allows for lower teams to have advantage over some higher teams in the second round, which means a team with a lower seed, and in term lower skill, can have more rewards than a higher skilled team (Baumann, Matheson, Howe, 2010, 12).

Experimental Design/Research Plan Goals:

The independent variables included the historical ringer percent gathered from previous tournaments, a new predicted ringer percent based on performance that would be calculated during a tournament, and the type of tournament to be played. The ringer percentages will be further explained later but the type of tournament played will either be double elimination or Horseshoe Mania. The dependent variables would be the fairness of the tournament itself and the fairness of the games within the tournament. Fairness would be defined by minimizing the differences in average skill between all teams. The controls would be the rules and objective of the horseshoe game and the material inside the horseshoe pits. The rules would be the NHPA standard rules, and the objective would be having the most points after each player throws 10 frames or 20 shoes. The material inside the horseshoe pits will be clay.

The experiment that will be carried out requires access to historical ringer percent data and an IDE to help writing code. It was noted earlier that the ringer percent data would be explained. Here is that explanation. The ringer percent is defined as the total number of ringers thrown in one session of horseshoes divided by the total number of horseshoes thrown. The historical data would be gathered from an online database that records weekly averages in the CCHC (Sluys). The new ringer percent would be calculated by the algorithm by averaging two quantities. The first value (the actual ringer percent) would be the amount of ringers divided by the total amount of horseshoes thrown. The second value (the theoretical ringer percent) would be the amount of shoes that almost were ringers divided by the total number of shoes. This way, the ringer percent takes into account an actual amount of ringers and how many ringers the person would have received had they played optimally, which would help take into account if the person started throwing better (having the actual ringer percent become closer to the theoretical ringer percent).

Once the algorithm has a working prototype, it will first be tested in a trial tournament then later applied to a real tournament. To test the algorithm, the ringer percent data will be retrieved from an excel sheet created from the historical ringer percent data. There will also be new ringer percent data gathered at the beginning, created from having each player not in the database throw 20 shoes and following the process described in the paragraph above. These percentages will be used in the algorithm to rank players into 4 different classes (A, B, C, and D) by seeding each player similarly to how professional sports seed teams (Schwenk, 2000, p. 140). The algorithm will then match A Class players with D Class players, and B class players with C Class players. As the tournaments go on, new ringer percent data will be brought in, and reseeding will be done for the teams similar to the NFL playoffs (Baumann et. al, 2010, p.10). The algorithm will be tested in different double elimination tournaments and Horseshoe Mania tournaments. It could also be tested with an increase of outliers in the set of all player ringer percentages.

Risk/Safety Concerns:.

During the game of horseshoes, people will throw metal horseshoes at high velocities. As such, it may be unsafe to be near horseshoe pits, as the horseshoes may bounce unpredictably and could possibly hit someone.

To address this safety concern, 4' tall fences recommended by the NHPA are put in place to prevent the horseshoe from hitting people (NHPA, 2019, p. 41). In general to help fix this issue, just make people aware of it and try to avoid it. For example, people should not move in front of a horseshoe while it's in the air, or try to move away from any horseshoe headed in their direction as quickly as possible.

Data Analysis:

To analyze the data from the CCHC, I will have the program look at the average ringer percentages for the week and rank people accordingly. These ranks will be set up in classes, A Class, B Class, C Class, and D Class; where A Class has the most skill and D Class has the least skill. To keep each tournament different from the next, a person from A Class will be partnered with a person from D Class every time, and a person from B Class will be partnered with someone from C Class, but the partners won't be the same every tournament.

However, this process might not be entirely accurate if a person had games of varying skill levels. For example, in game 1 a person might throw a 60% average and game 2 that same person might be "off" (throw much worse, say 15%). So, if their average score for that week is an outlier compared to their normal average scores, it would not make sense to update their rank drastically due to a difference in one week. In this case, the program would rely more heavily on previous weeks before and after the outlier week.

During a test tournament, as the tournament goes on data would be gathered at the end of each round about the score of the game, how many shoes were thrown, and how many ringers were thrown by each person. I could then use the Markov Model to estimate the win rate for each team and reseed based on this (Khatibi A. et al., 2015).

Potential Roadblocks:

One potential roadblock is gathering people for a horseshoe tournament in the CCHC because of COVID-19. The safety of the horseshoe players would be paramount, so COVID safety guidelines would be enforced.

Another potential roadblock could be corruption of the excel file that stores my data. A solution would be to make multiple backups, for example two backups could be made per week, or one backup per day and another weekly backup.

References:

NHPA. (2019, January 1). *The Official Rules/Regulations, Guidelines, Specifications*. Horseshoepitching.com. https://www.horseshoepitching.com/wp-content/uploads/2018/03/RGS2019Web.pdf

Baumann, R., Matheson, V. A., & Howe, C. A. (2010). Anomalies in Tournament Design: The Madness of March Madness. *Journal of Quantitative Analysis in Sports*, *6*(2). https://doi.org/10.2202/1559-0410.1233

Khatibi, A., King, D. M., & Jacobson, S. H. (2015). Modeling the winning seed distribution of the NCAA Division I men's basketball tournament. *Omega*, *50*, 141–148. https://doi.org/10.1016/j.omega.2014.08.004

Schwenk, A. J. (2000). What Is the Correct Way to Seed a Knockout Tournament? *The American Mathematical Monthly*, *107*(2), 140–150. https://doi.org/10.1080/00029890.2000.12005171
(Note: This link doesn't grant full access, but if the reader searches in Google Scholar "NCAA tournament seed", and clicks the PDF, they gain full access.)

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *The American Economic Review, 83*(5), 1281-1302. Retrieved October 12, 2020, from http://www.jstor.org/stable/2117561

Sluys, C. (n.d.). *Https://cchc.club/*. Central CT Horseshoe Club. Retrieved October 29, 2020, from https://cchc.club/

Timeline:

My project will be divided into 4 phases: Brainstorming and Researching, Data and Prep, Testing the prototypes, and Perfecting one prototype.

Phase 1 is brainstorming and researching. This was already done prior to starting this project. During this phase, it was decided what project will be done, and background research was completed.

Phase 2 is gathering historical data and preparing to test my algorithm in the future. This includes creating prototypes of the algorithm, using the CCHC website to gather historical ringer percent data, and making contact with the CCHC to set up a tournament.

Phase 3 is testing the prototypes. First, I would set up dummy tournaments to see what I could improve on the prototypes. The tests will search for how many games are deemed unfair using criteria. The prototype would then be altered to reduce this number. Once the prototypes are deemed ready to test on an actual tournament, a tournament at the CCHC will be set up and the prototypes would be tested there.

Phase 4 is perfecting one prototype. At this point, the prototype that leads to the least amount of unfair games and deemed the best will be chosen. The UI would be polished, the program would be made dummy-proof (test all possible cases), and I could test how people use the program to help debug any issues.