Mean Value of Red-Blue-Green Hackenbush Trees

Jacob Sampson

2017

Abstract

Hackenbush is one of the most visual demonstrations of the link between surreal numbers and their arithmetic and combinatorial games. Addition for the case of stalks, and the more general hackenbush trees, often doesn't need the translation to surreals to be computed. This paper develops an algorithm for addition and multiplication on RBG hackenbush stalks, and shows how trees can be simplified to stalks. Most RBG trees have values that are not surreal numbers, but have an invariant called the mean value that is surreal. I prove a theorem about the mean value of RBG hackenbush trees, and show how a player can compute their best strategy.

Contents

Chapter 1

Introduction

Hackenbush is a game invented by John Conway and is often used to introduce the connection between combinatorial games and surreal numbers. Hackenbush is a convenient way to show how adding two games is equivalent to adding surreal numbers and other game values. This is often done by finding the value of the each individual Hackenbush graph and then summing those values. For the game in figure 1., this method is probably best.

For simple stalks, as seen in figure 1.2, this is often an unnecessary process. The game labeled (d) has one blue edge and one red edge, so neither player has an advantage over the other. The oppositely colored edges cancel each other out and the overall game equals zero. This paper develops an algorithm for adding Hackenbush stalks without deriving their individual surreal values. Given two stalks, this method will give a single stalk whose value is the sum of the values of the inputs. Now there is a visual idea of what happens when two hackenbush games are added. This method is naturally applicable to Hackenbush trees and can be modified to find Red-Blue-Green Hackenbush sums. Addition is quickly followed by subtraction because negation is easy in combinatorial games. Multiplication is more difficult and not as easily applicable as addition, but it further builds up an arithmetic on Hackenbush stalks with surreal values. The final result is a theorem on mean values for Red-Blue-Green Hackenbush trees, whose values are not surreal numbers. The theorem also includes an alternate way to calculate the mean value of a Hackenbush tree. This calculation is much more inefficient than the other method, but the process shows players to the optimal ways they should play on a RBG hackenbush tree, to perhaps do better on than the mean value.

Given any arbitrary hackenbush game with a fairly complicated graph, it is often difficult to find the value, even with computers. Finding the optimal strategy for each player can be a complex problem that is not always obvious. Computing the value of an arbitrary game is NP-Hard. Many results and current research on Hackenbush typically focus on a specific type of Hackenbush graphs to find clever ways to figure out values of complicated Hackenbush positions. This paper focuses on trees.

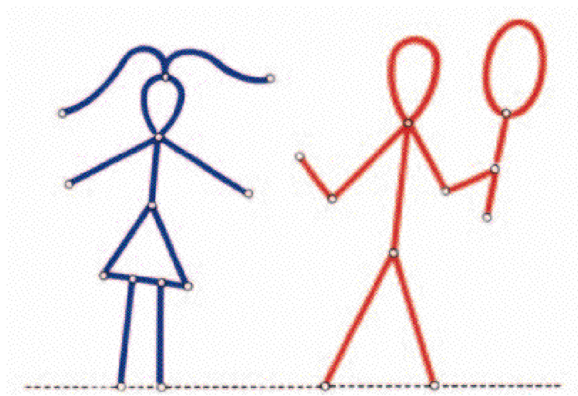

Figure 1.1:

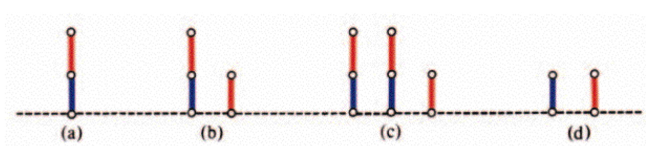

Figure 1.2: Hackenbush stalks

Chapter 2

Background

Hackenbush is a two-player game where the players alternate removing edges from a graph or collection of graphs. Typically the players are called **Left** and **Right**. The edges are all colored red, blue, or green, and one vertex is regarded as the ground, which is represented as a line at the bottom. All edges have a path to the ground. The left player can only cut blue edges and the right player can only cut red edges. Either player can cut a green edge. After each move, edges that are not connected to the ground are no longer in play. For example, examine the boy in figure 1.1. If right removes the boy's hand, the rest of the ballon will no longer be available to right after that move. The loser in this game is the player who has no more moves to make on their turn. For simplicity, this paper refers to a hackenbush game with only red and blue edges as RB Hackenbush, and hackenbush games that include green edges as RBG Hackenbush.

Hackenbush is a common example of a combinatorial game. Most texts define a combinatorial game similiarly to Richard Nowakowski, from his book **Games of No chance**:

1. There are two players moving alternately;
2. There are no chance devices and both players have perfect information;
3. The rules are such that the game must eventually end; and
4. There are no draws, and the winner is determined by who moves last.

The study of the structure of combinatorial games, and the methods used to figure out the outcome of a game, are the focus of a branch of mathematics called Combinatorial Game Theory. While humans have been playing and studying games for hundreds if not thousands of years, the modern study of Combinatorial Game Theory began with John Conway. He invented a class of numbers, called the surreal numbers, that are all values of certain combinatorial games. Not all games are surreal numbers, however. Hackenbush is the standard for introducing the connection between surreal numbers and other game values and outcomes of certain hackenbush games.

The term ruleset refers to the system of playable rules. The ruleset of chess would be the description of the board and pieces and how they move and interact. A game of chess would be a certain position of pieces on the board.

Definition: A **game** is an individual position of a combinatorial game.

When left and right make **moves** on a game, they change it to a different game.

Definition: Let G and H be games. H is a **Left Option** of G if Left can move from G to H in one move. H is a **Right Option** of G if Right can move from G to H in one move.

The formal notation of a game G's options is ordered pair notated $\{G^L|G^R\}$. The set G^L contains every left option of G and the set G^R contains every right option of G.

All games can be partitioned into one of four **Outcome Classes**. They are

- $\mathcal{N}$: first player (the Next player) will win
- $\mathcal{P}$: second player (the Previous player) will win
- $\mathcal{L}$: Left will win
- $\mathcal{R}$: Right will win

These outcomes are under the presumption that both players are playing rationally and optimally.

The outcome class of the game in figure 1.1 is not too complicated to work out. Left has 14 edges available and right has only 11. No player can change the number of edges available to the other. If left plays rationally, he will remove edges starting from the top and going down, careful to only remove one edge at a time. Eventually right will run out of edges before left does, so left can always win this game.

A natural question to ask is: by how much? If left moves first and both players play rationally, right can only make 11 moves after each of left's 11 moves. Then left will cut her 12th edge, and on right's turn he has no moves and loses. At the end of this game, left still has two moves to make before she loses. This is the beginning of the intuition behind the values of games, and the equivalence between games and numbers.

The simplest position of hackenbush is a graph with no edges at all, shown in figure 2.1. Left and right both have no options to move anywhere, and the outcome class is $\mathcal{P}$. This game is called 0, or the empty game. Notice that both options are empty, so the notation is $0 = \{|\}$.

A single blue edge is a one-move advantage for left, so this game is in the outcome class $\mathcal{L}$. Blue edges are considered positive, so a single blue edge is given the value 1. Notice that left moves to zero while right has no moves, so $1 = \{0|\}$.

Similarly, a single red edge is $-1 = \{|0\}$. Notice that in figure 1.2(d), there is one blue edge and one red edge. This is a second player win. The outcome class is $\mathcal{P}$, just like the empty game. A one-move advantage for left and a one-move

Figure 2.1: The Empty Game

advantage for right cancel each other out. This is an informal example of why $1 + (-1) = 0$ in hackenbush stalks.

Red-Blue hackenbush stalks can be used to represent the surreal numbers. A **surreal number** is an ordered pair of sets of previously created surreal numbers. The sets are known as the left set and the right set. No member of the right set may be less than or equal to any member of the left set. When green edges are introduced, they violate the conditions of surreal numbers. It is still possible to study these game values, and they do form a partial ordering with the surreal numbers.

The method used to create these numbers hints at a hierarchy in the surreals. The first surreal number is 0, or $//$. Next, the numbers 1 and -1 can be constructed, as $\{0|\}$ and $\{|0\}$, respectively. The following surreal numbers that can be created are $2 = \{0, 1|\}, 1/2 = \{0|1\}, -2 = \{|0, -1\}, -1/2 = \{-1|0\}$. Because 0 was constructed before the other numbers, it is said to be simpler than all the other surreals. Similarly, 1 is simpler than 2 and 1/2 is simpler than 3. The method to find the simplest number between two values is called **Simplicity Rule**. This is finding either the smallest integer between the two, or else the fraction between them having the highest power of two in the denominator. In this way simplest means the surreal number constructed earliest.

The figure below shows the beginning of the surreals and their 'birthday', the order in which they can be constructed.

A single green edge is notated $*$, pronounced star. Both left and right can remove a green edge, so $* = \{0|0\}$. The outcome of $*$ is a first-player win, so $* \neq 0$. Surprisingly, $* + * = 0$, and it can be shown that $*$ is less than any positive surreal number, and greater than any negative number. For these reasons, it is often said that star is confused with zero. In the class of game values they are incomparable, but they have some similar characteristics.

Even though RBG hackenbush positions are not numbers, it is still possible to define how much a move on any game is worth.

Definition: A **Left incentive** of a game G is denoted $\Delta^L(G)$, and is a game of the form $G^L - G$. A **Right incentive** of a game G is denoted $\Delta^R(G)$, and is a game of the form $G - G^R$.

The following are some results from Siegal, which will be used in the next section.

Proposition 1: If x is a number, x^L, x^R are any members of the left and right set of x, then $x^L < x < x^R$. In particular, every incentive of x is negative.

Proposition 2: If G is not a number, then G has both a left incentive $\Delta^L(G)$ and a right incentive $\Delta^R(G)$ such that

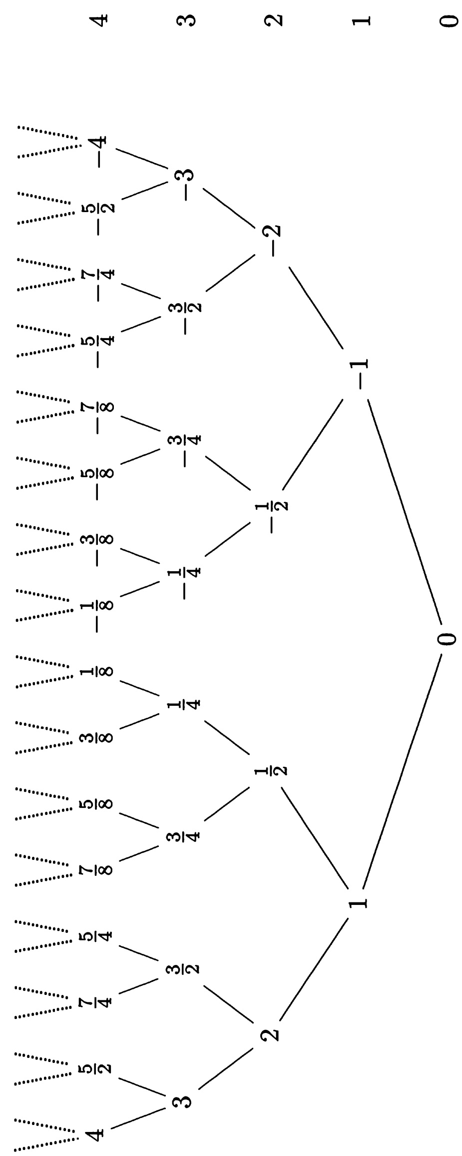

Figure 3.1. The number tree (with birthdays labeled on the right).

Figure 2.2:

$\Delta^L(G) \geq -x$ and $\Delta^R(G) \geq -x$ for every number $x > 0$.

Number Avoidance Theorem: Suppose that x is equal to a number and G is not. If Left (resp. Right) has a winning move on $G + x$, then he has a winning move of the form $G^L + x$ (resp. $G^R + x$).

Proof: Suppose left has a winning move of the form $G + x^L$. Then certainly $G + x^L \geq 0$¿ But G is not equal to a number, so $G + x^L > 0$. Therefore left has a winning move on $G + x^L$. By induction on x, this move has the form $G^L + x^L$, so that $G^L + x^L \geq 0$. Since x is a number, $x^L < x$, and in fact $G^L + x > G^L + x^L \geq 0$.

If G is a hackenbush game, and not a number, eventually all green edges will be cut and a subposition of G will be a number x. Assuming left or right goes first, and both play optimally, each player can guarantee an invariant on G, named the left stop and right stop.

Definition: The Left stop $L(G)$ and the right stop $R(G)$ are defined recursively by

$$L(G) = \begin{cases} G & \text{If G is equal to a number;} \\ max(R(G^L)) & \text{Otherwise;} \end{cases}$$
$$R(G) = \begin{cases} G & \text{If G is equal to a number;} \\ min(L(G^R)) & \text{Otherwise.} \end{cases}$$

Suppose Left and Right play G with left moving first. The number x that they reach, assuming both play rationally, is $L(G)$. Likewise, $R(G)$ is the number reached if right moves first.

The mean of a game G is another invariant. It is a number $m(G)$ that approximates many copies of G, so that $nG - nm(G)$ is bounded by a constant independent of n. The following theorem has been proven independently several times and in different. This paper uses the form used in Siegal.

Mean Value Theorem: For all games G,

$$lim_{n\to\infty} \frac{L(n\times G)}{n} = lim_{n\to\infty} \frac{R(n\times G)}{n},$$

and in particular both limits exist.

Definition: The mean, or **mean value** of G is given by

$$m(G) = lim_{n\to\infty} \frac{L(n\times G)}{n}.$$

Now there is enough background to state my results and start sketching the proof.

Chapter 3

My results

The main result that has motivated this paper is stated below. This section will outline the main ideas of the proof and then go through the details.

Theorem: The mean value of any RBG hackenbush tree is the simplest number between the tightest bound obtained by replacing green edges with red and blue edges.

It is easier to prove that a similar result holds for a particular subset of RBG stalks instead of trees, which will outline a method of proof that will be repeated throughout this section. Using propositions from other sources, the method of proof can be generalized to all RBG stalks and finally RBG trees.

The particular type of RBG stalks are those that have exactly one green edge at the end of the stalk, or with the furthest path from the ground.

First, it is important to remember that any player has an incentive to remove these top green edges when available. If k is a number, by Proposition 1 every incentive of k is negative. Suppose G is a hackenbush game with red-blue edges and one green edge at the top. From Proposition 2, G has right and left incentives that exceed every negative number. Since G has one green edge at the top, it is not a surreal number. The red-blue stalk however, must be a number. That number will be denoted x from now on.

Lemma 1: the best move for one player on a RBG stalk that has exactly one green edge at the top is to remove the green edge.

Proof: Let x denote the number that is the value of the red-blue stalk of G, or the number G is infinitesimally close to. If $x = 0$, then G is just one green edge, and the only move is to cut the green edge.

If $x < 0$, then right will have an inherit advantage in this game. To see why right would have the most advantage by removing the green edge, look at how incentives are calculated. The equation for right incentives are $\Delta^R(G) = G - G^R$, so when right removes a red edge, G^R is strictly less than G. This is because the right incentives for G are strictly negative for any move on the red edges, and confused with zero for the move cutting the green edge. When right moves on the green edge, $G^R = x$, so G^R is confused with G and the incentive is confused with zero. Thus red's best move is to remove the green edge.

If $x > 0$, then left will have the most advantage by removing the green edge. The argument is similar to the case for $x < 0$, but replacing right with left, red with blue, and the equation for incentives to $\Delta^L(G) = G^L - G$. □

If the stalk is negative like in the proof, it is important to remember that right has a winning move by removing the bottom edge for this game, and leaving left with no moves. The best move is not necessary the winning move, because the best move will leave the greatest advantage to right in the context of other games. This idea will be important in the proof of the mean value of RBG stalks.

Lemma 2: The mean value of a RBG stalk with one green edge on top is the number obtained from removing the green edge.

Proof: The mean value of a game G is defined as $\lim_{n\to\infty} \frac{L(n\times G)}{n}$, where $L(n \times G)$ is the left stop of n copies of G. The left stop x is a surreal number, and can be thought of as the final score after left and right alternate playing on $n \times G$, with left moving first. The game $n \times G$ is n copies of G. Assume the value of the RB stalk is positive, or equivalently that $x > 0$. So when left starts, she will remove any green edge, from the first lemma. Right will make a move, and will want to preserve as many moves as possible, so will cut another green edge. On left's next move, he will remove another green edge, because of the number avoidance theorem. Eventually all green edges will be cut, and the left stop is the number represented by the RB stalk times n. This process holds for all values of n, so the mean value is x, the number obtained from removing the green edge. □

This can be extended to any RBG stalk. The proof is similar to the one above so won't be fully written out. A concept from the book **Winning Ways**, an early introduction to Combinatorial Game Theory will be needed in the argument. It is reworded below to better fit the subset of Hackenbush positions this paper deals with.

Corollary 1: No sane person will chop an edge beneath the green edge that is closest to the ground while there's any other edge to chop.

So to prove the mean value of a RBG stalk, H, is the simplest number between the tightest bound obtained by replacing green edges with red and blue edges, it is enough to prove that the left stop of n copies of H is n times whatever surreal number forms the "base" below any green edges. The simplest number between the tightest bound when green edges are replaced by red and blue is that surreal number. This fact is trivial, because any stalk with fewer edges will be simpler than a stalk with more edges.

lemma 3: The mean value of any RBG stalk, H, is the simplest number between the tightest bound obtained by replacing green edges with red and blue edges.

The argument for the left stop of $n \times H = n \times x$ follows the same pattern as the proof of lemma 2 above. When playing on n copies of H, both players will remove an edge above or on the lowest green one for any stalk, by corollary 1. If one copy of H is reduced to x, the players will move to other copies by the number avoidance theorem. Eventually there will be a game that is $n \times x$, so the left stop of H is x.

The final theorem uses the concept from corollary 1 in the last proof, but it should be rewritten to fit the tree terminology.

Corollary 2: No sane person will chop an edge in the Red-Blue subtree (?) while there's another edge to chop.

This corollary will be used in the same way as in the proof of the mean value of RBG stalks.

As hinted above, the mean value will be the value of the sub-tree of red-blue edges. That is, the largest graph started from the ground and stopping, on all branches, just before the first green edge on all branches. The proof will show that the value of the sub-tree is indeed the mean value, and also the simplest number between the tightest bound obtained by replacing green edges with red and blue edges.

Here is the theorem restated followed by a formal proof:

Mean Value Theorem For RBG Hackenbush Trees: The mean value of any RBG hackenbush tree is the simplest number between the tightest bound obtained by replacing green edges with red and blue edges.

Proof:

Let G be an arbitrary RBG hackenbush tree, and x the number that is the value of the Red-Blue subtree. If the left stop (or right stop) for n copies of G is $n \times x$ for any positive integer n, then the mean value of G is x, by definition.

Suppose there are n copies of G and $G > 0$. By corollary 2, left will move on some edge not in the red-blue subtree x of some copy of G. Right will move the same way for the same reason. Eventually, a copy of G will be reduced to x. Left and right both have negative incentives if they move on x. If either player cuts an edge above a green on any particular branch of another copy of G, their incentive will be confused with zero, but still greater than any negative number. This process continues until all copies of G are reduced to x, giving $n \times x$ as the left stop for any n. Therefore, x is the mean value of G.

Now all there is to show is that the mean value is the simplest number between the tightest bound obtained by replacing green edges with red and blue edges. Take the green edges that separate x from the rest of game G. These edges will start subtrees on G on multiple branches of G. Consider one of these branches. The simplest number obtained by replacing the green edges with blue or red will be the value of the red-blue branch before the first green edge. This idea was shown in a previous proof. The branch is shorter, so will necessarilly be the simplest number. Since the simplest number within the bounds for each branch is the value of that branch, applying the replacement algorithm across the whole tree will give the desired upper and lower bounds that have the mean value as the simplest number between them. $\square$

The tightest bound computation may seem redundant, but it will be necessary for a player (or computer) to find the values that are possible when choosing which green edges to cut. The mean value is not necessarily the actual value that will happen for an optimal strategy, and in the context of other games, left or right may want to cut lower or higher green edges to force the other player to move a certain way. It can also simplify certain calculations. For example, take a game that is two different RBG trees, one with 3 green edges and the other with more than 3. Based on the coloring of the trees, right may have less incentive to move on the tree with 3 green for at least 3 moves. Left could claim these edges, essentially "coloring" them blue, and giving that tree a number value. Numbers are easier to work with computationally than other game values found in hackenbush.

Chapter 4

References

- Berlekamp, E. R., J. H. Conway, and R. K. Guy. Winning Ways. London: Academic, 1981. Print.
- Siegel, Aaron N. Combinatorial Game Theory. Providence: American Mathematical Society, 2013. Print.
- 10, 2009 February. Surreal Numbers and Games (n.d.): n. pag. Web.
- Knuth, Donald Ervin. Surreal Numbers: How Two Ex-students Turned on to Pure Mathematics and Found Total Happiness; a Mathematical Novelette. Reading, MA: Addison-Wesley, 2011. Print.
- Conway, John H., and Richard K. Guy. The Book of Numbers. New York, NY: Copernicus, 1996. Print.
- Conway, John Horton. On Numbers and Games. N.p.: Blackwell Science, 1976. Print.
- Albert, Michael. Lessons in Play: An Introduction to Combinatorial Game Theory. S.l.: AK PETERS, 2017. Print.
- Davis, Tom. ”Hackenbush.” (2011): n. pag. Geometer. 15 Dec. 2011. Web.