Greedy Tree Simulation

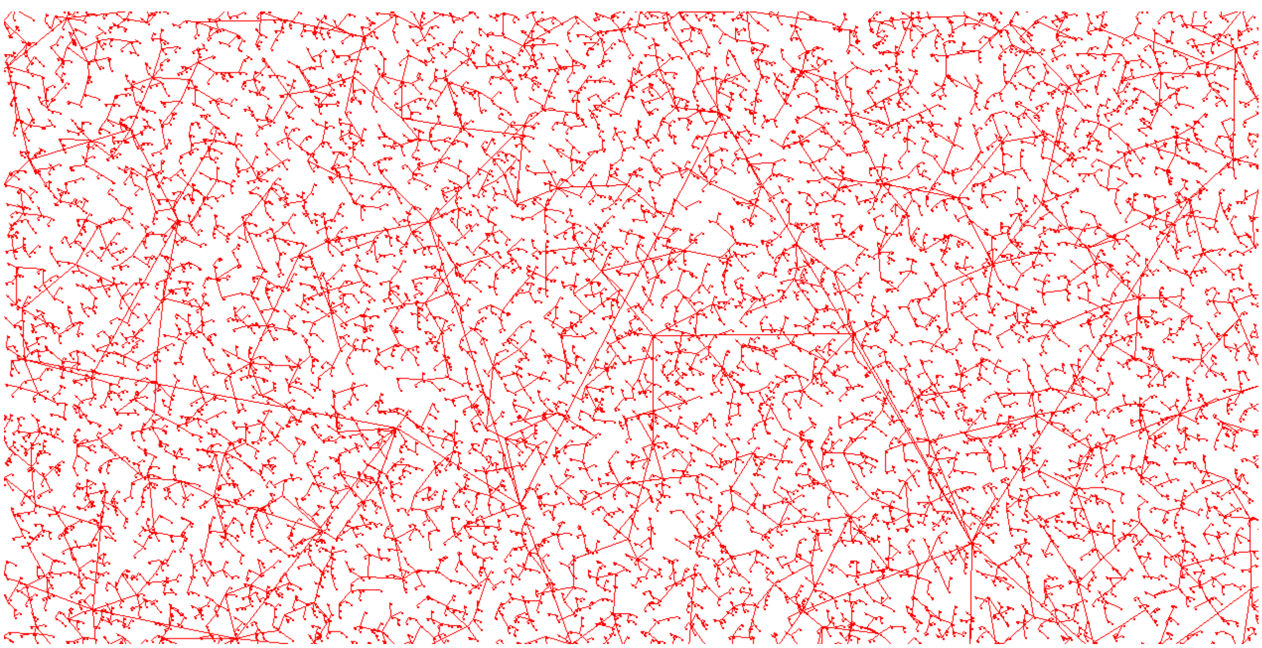
Weijian (James) Han

Description

This project is a simulation of the greedy tree, in which a given number of points are randomly placed on a grid with predetermined dimensions (but is supposed to be a representation of an infinite plane), and each point is connected to the closest of the previous points. A number of analyses are done with this simulation, including zooming in on parts of the greedy tree to examine its resemblance with the entire tree, dividing the tree into distinct regions extending from each of the source points, and calculating the fractal dimension of the boundaries between regions.

Part I: Drawing the Tree

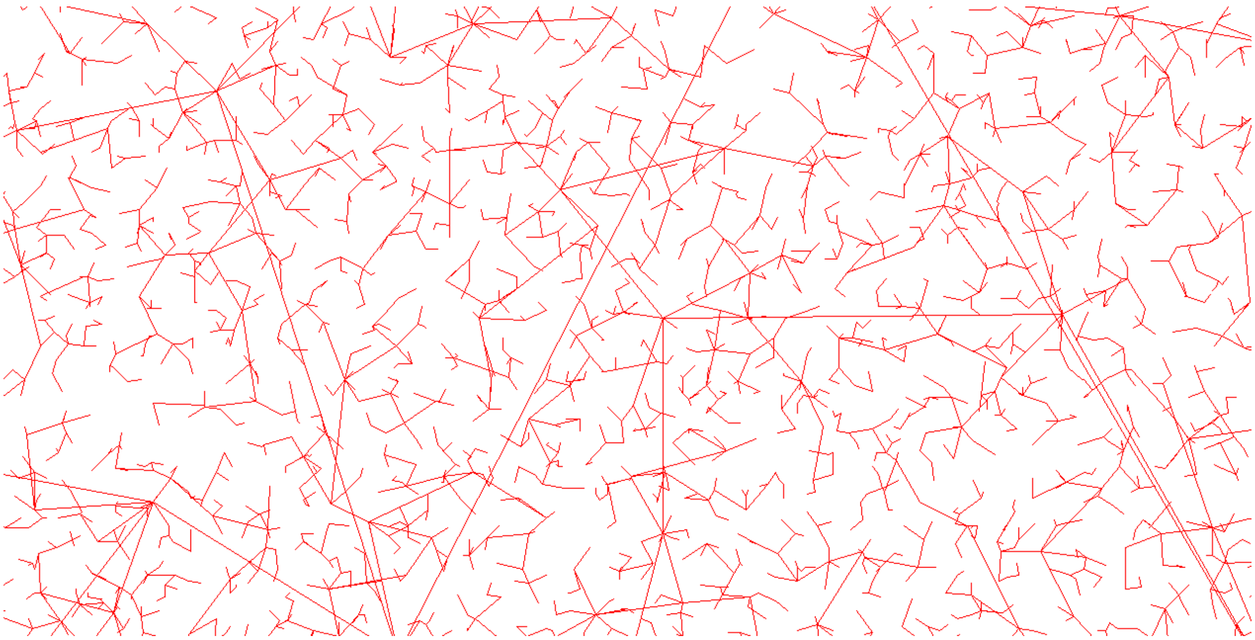
Before any analyses can be done, a greedy tree has to be drawn, and this process is the same for all parts of this project. First, the appropriate parameters are chosen, including the number of points and the grid size. The grid size is fixed to be 1280 by 680 pixels, and the number of points vary is from simulation to simulation based on the speed and memory requirements of each simulation. Then, the points are randomly placed on the grid one by one, and a line is drawn to connect each point to the closest previous point. This process is repeated until all of the points are placed on the grid. (Figure 1)



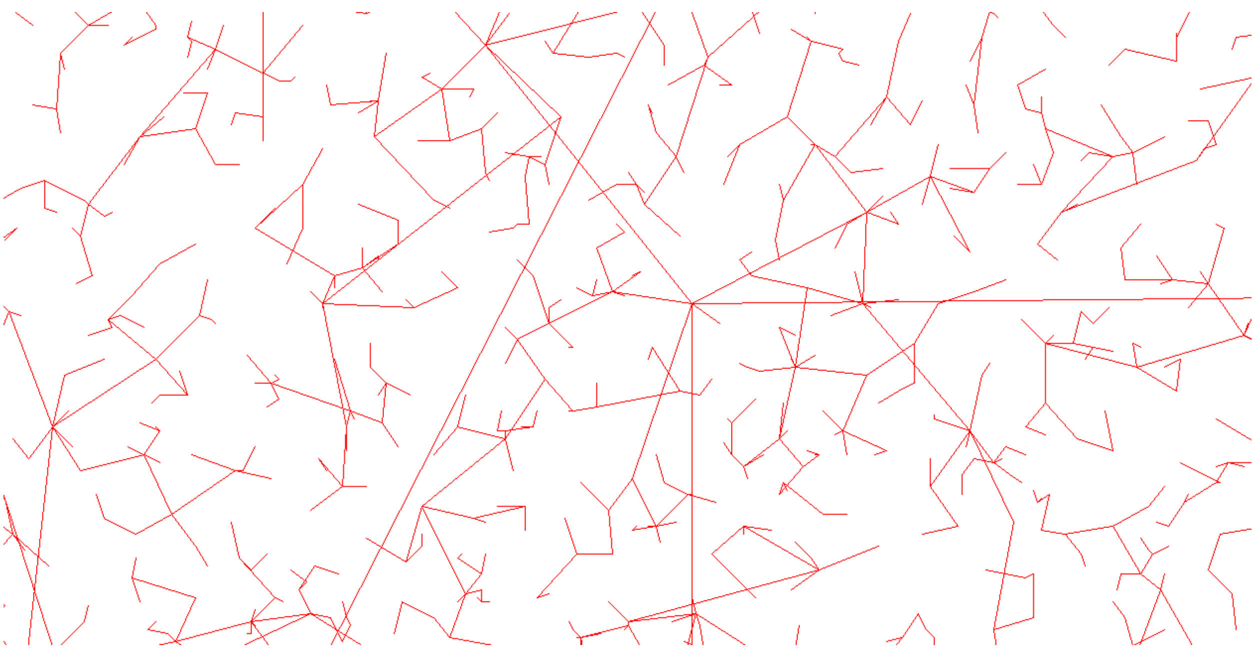
(Figure 1: Original Greedy Tree)

Part II: Zooming In

The greedy tree is an example of a fractal, exhibiting self-similar patterns, meaning if you zoom in at a piece of the greedy tree, it will look just like the original picture. This self-similarity is evident when zoomed in at a half (Figure 2) and a quarter (Figure 3) the scale of the original picture, but since we can only have a finite number of points in this simulation, this pattern becomes not so obvious when zoomed in further.



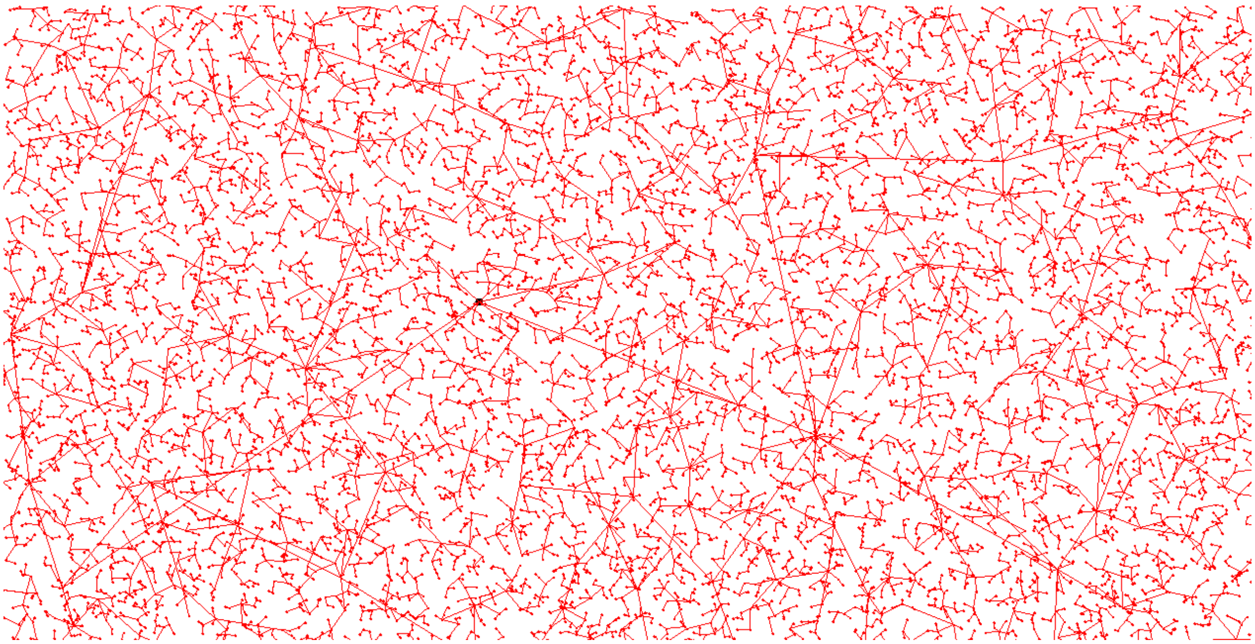
(Figure 2: Zoomed In at Half Size)



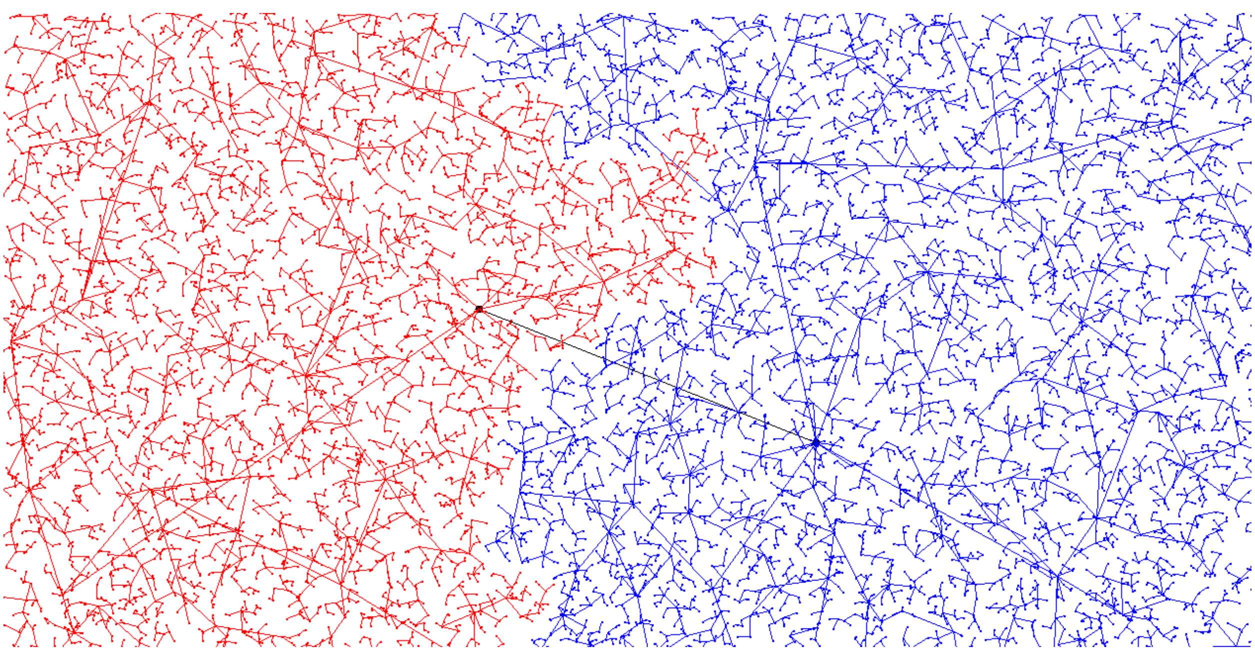
(Figure 3: Zoomed In at Quarter Size)

Part III: Dividing the Tree

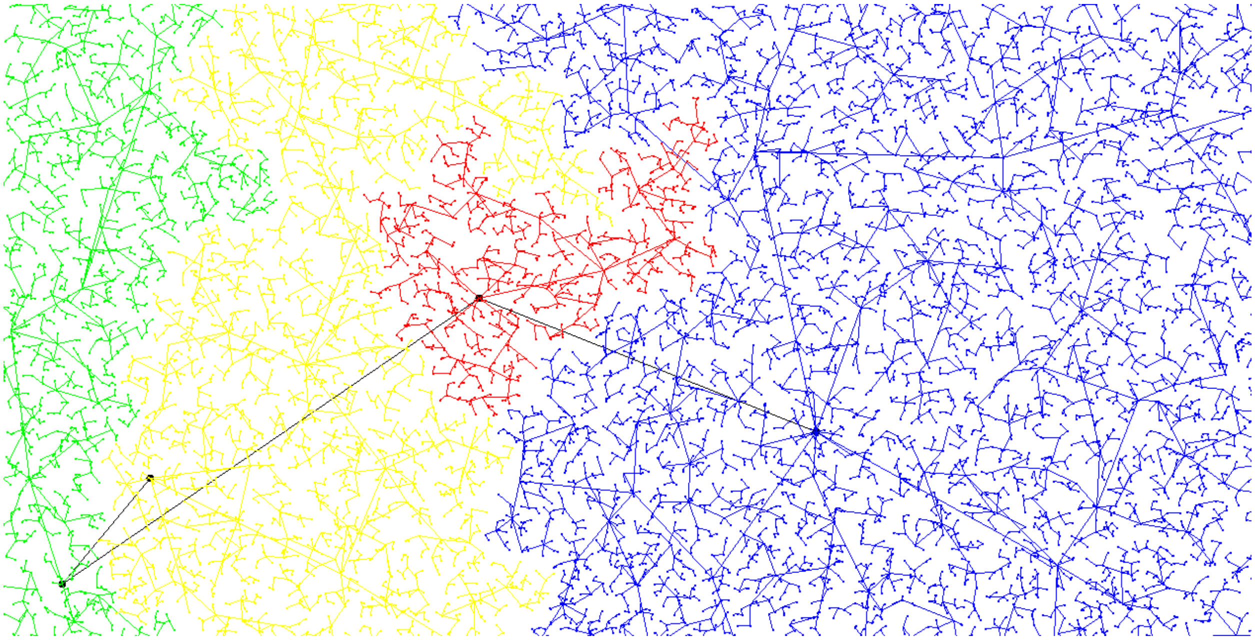
In this simulation, the first few points act as source points, where all other points will connect from. Each source point is given a color, and since every future point has to connect to one and only one other point, it is given the same color as the point it connected to, and each point is ultimately connected to one of the source points, so all points originating from the same source point shares the same color. The following pictures show the same 10,000 point greedy tree being divided up using 1, 2, 4, and 8 source points.



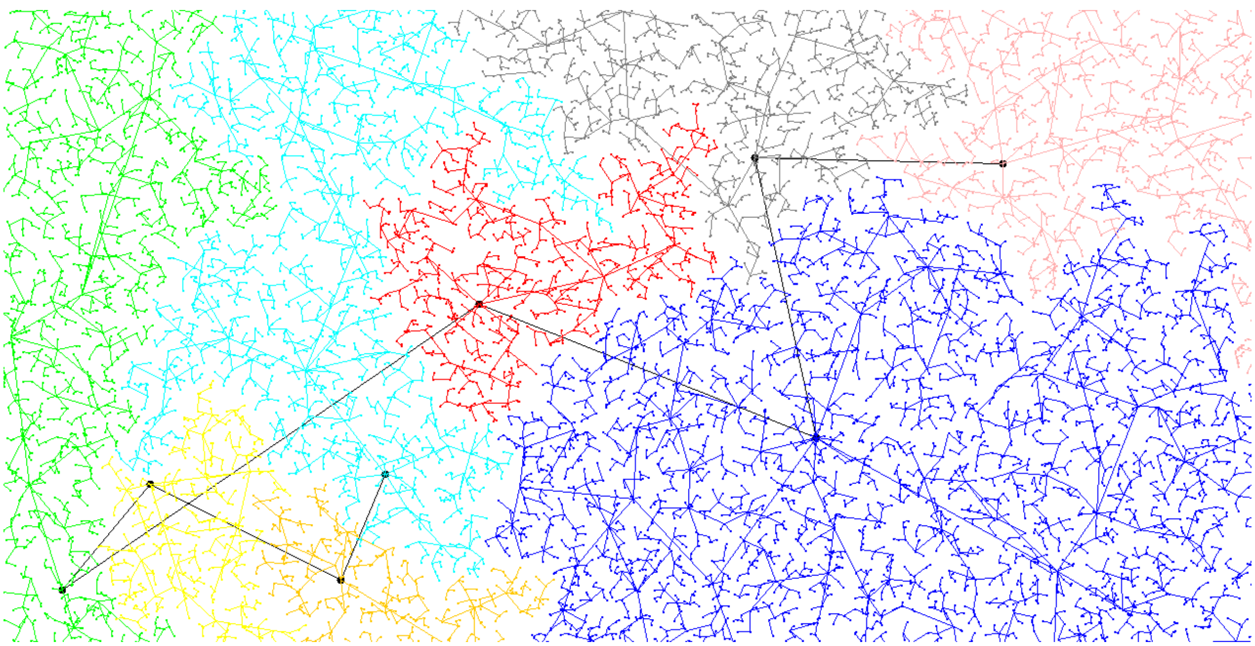
(Figure 4: One Source Point)



(Figure 5: Two Source Points)



(Figure 6: Four Source Points)

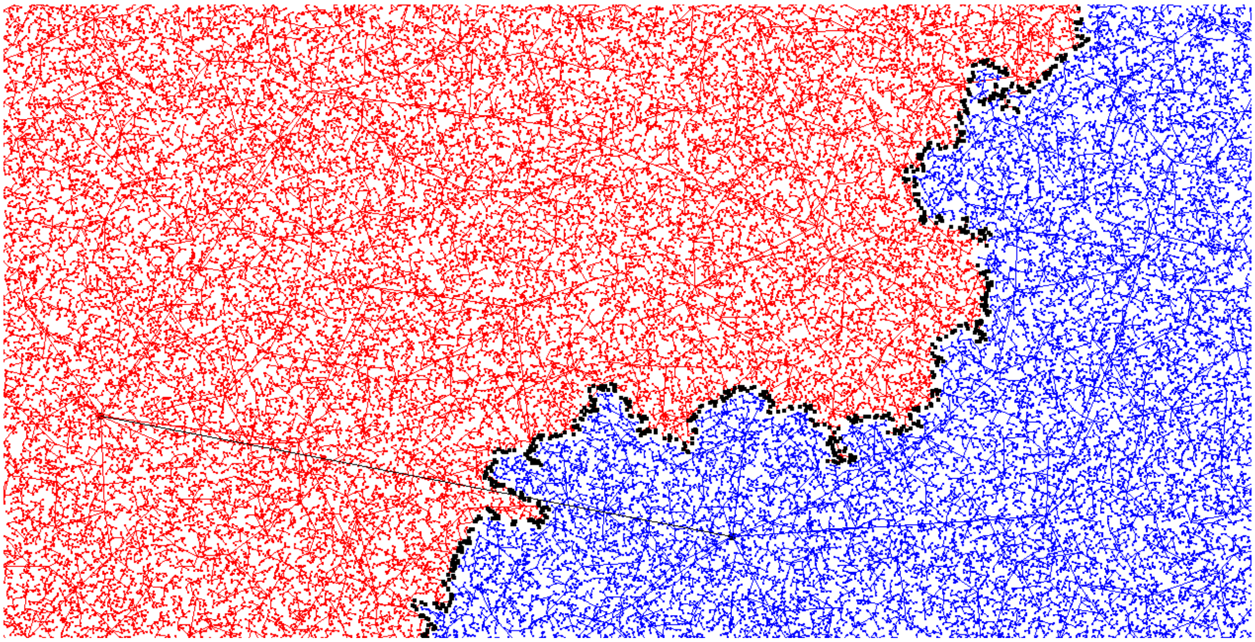


(Figure 7: Eight Source Points)

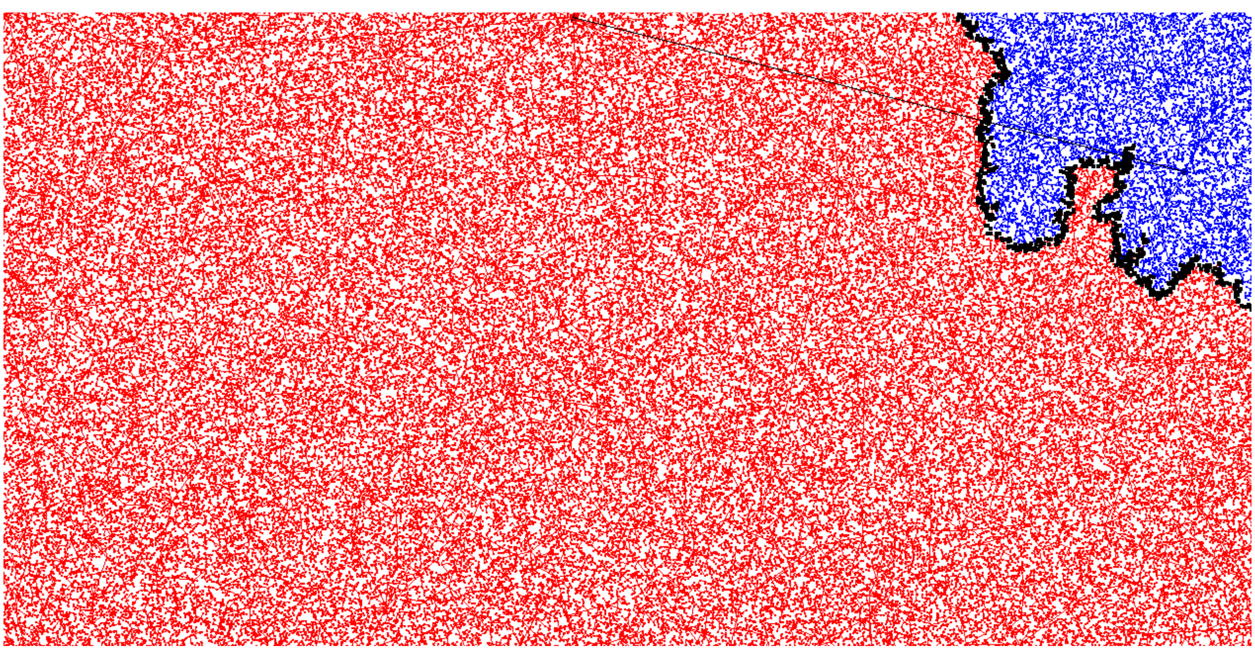
Notably, each of the eight regions in Figure 7 is a subset of one of the four regions in Figure 6, which itself is a subset of one of the two regions in Figure 5. This means that when cutting up the tree, only additional border lines are added, and the original border lines are kept in place.

Part IV: Counting Border Points

The first attempt to estimate the length of the border between two regions involved defining a set of “border points” which lie within a certain distance from the closest point of the other region. In the case of the following pictures, the red points within 15 pixels (arbitrarily chosen) from the closest blue point are considered border points. Although the number of border points can be easily counted, it is difficult to infer from it the length of the boundary.



(Figure 8: 3,644 Border Points out of 50,000)

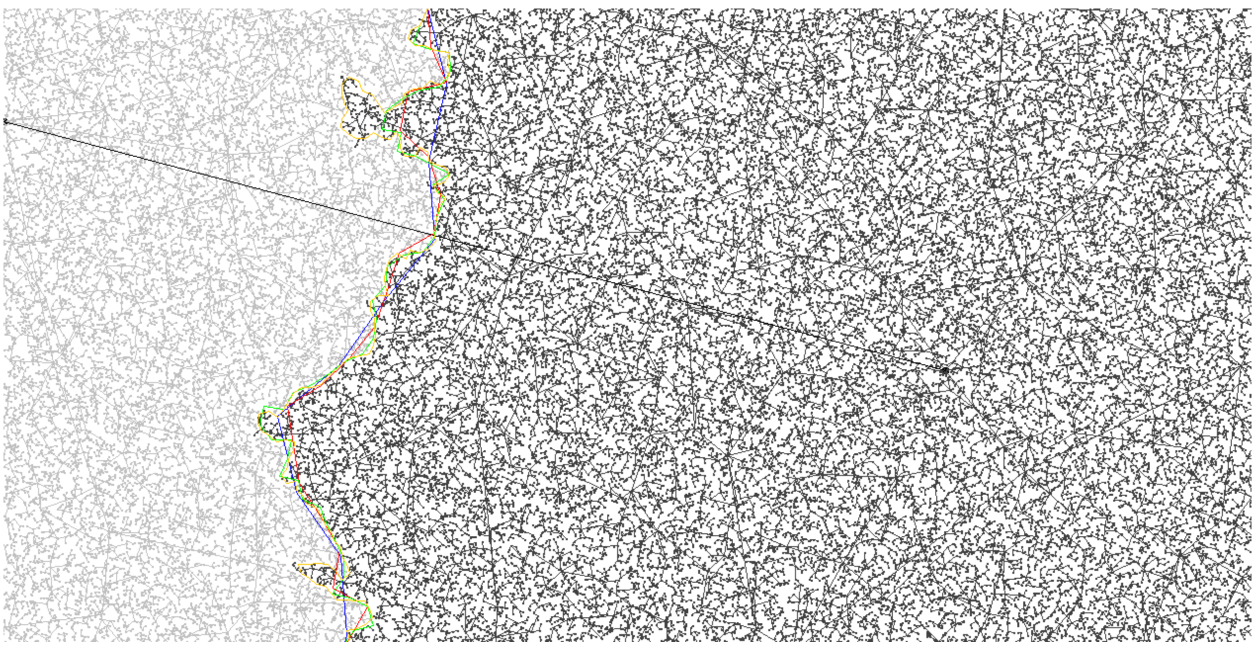


(Figure 8: 7,800 Border Points out of 100,000)

Part V: Estimating the Fractal Dimension of the Border

The second approach to estimating the "border length" involves drawing a line between the two regions, where each point on the line is equidistance to the closest point from either side of the border. However, it is impossible to find every point on this line, so we can estimate it by starting from one end of the border and finding the next point a distance *r* away that is equidistance from the closest points on each side until we reach the other end. These estimated points are then connected with line segments of length *r*, and the length of the border line can be estimated this way. By decreasing the size of *r*, we will get better results for the estimation, eventually converging to the true length of the border line between the two finite regions of the greedy tree.

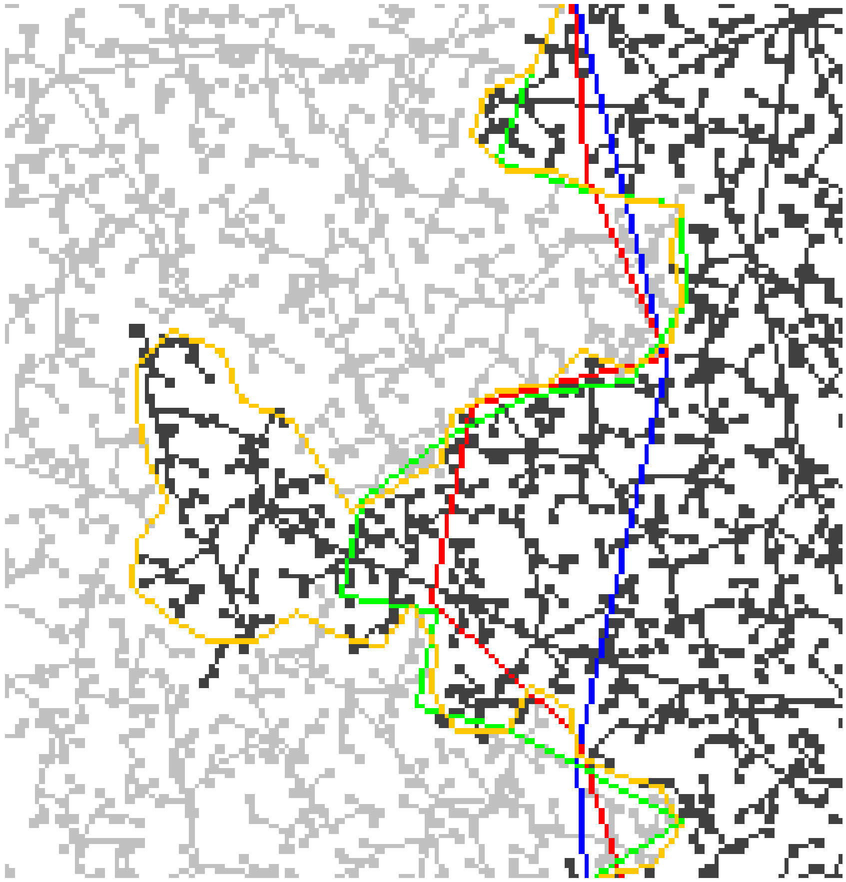
Since this simulation is a simplification of a greedy tree with infinite number of points on an infinite plane (therefore the border line would be of infinite length), it would make more sense to calculate the fractal dimension of the border rather than its length. A few simulations are done to draw border lines with different *r* values, and their corresponding estimates are recorded. As the number of points increase, these estimates begin to show a relationship between *r* and the estimated border length.



(Figure 9: Different Estimates of Border Length with 50,000 Points)

r	Estimated Border Length
80	1280
40	1640
20	2120
10	2790

Here is a zoomed in picture of a section of the border. The blue, red, green, and yellow lines correspond to *r* values of 80, 40, 20, and 10 respectively.



(Figure 10: Figure 9 Zoomed In)