Peter Rosenthal Homework 3 Writeup

April 02, 2021

Simulation of Flocks

For this homework set, unlike the previous ones, I opted to write my own code rather than using any of the starter codes. The reasons behind this were complicated, but it's related to my final project coming at the end of this semester. This ended up being a huge waste of time in the context of this homework assignment alone, but it certainly wasn't a waste of time overall in the end. The simulation was coded in TypeScript (a strongly typed superset of JavaScript) using primarily the creative coding framework p5.js creative boids like output, and then also my attempt at using it to dynamically draw graphs from scratch. The basic inspiration behind this all is that boids have always inherently bridged the gap between creative coding and biology/real sciences; the original paper was presented at SIGGRAPH and has been used in games, tv shows, and movies since then, but it's also the mother of all flocking algorithms studied in computational biology today. It was this connection that inspired me to try to merge creative coding and science in my own way.

Scaling Factors

The first part of the assignent asks us to investigate the various scaling factors of the simulation and how changing them can impact the result of the simulation. For all of the graphs in this section: simulations were run up to timestep 100, and time averages were taken accross all timesteps simulated. All data was averaged over the course of 20-100 realizations to reduce noise.

c_1

The first scaling factor is also called the attraction factor because it controls how strongly attracted the boids are to each other. At high attraction factors (relative to the repulsion factor, `c2`), the flock of boids collapses down into a dense cluster. At more moderate values of `c1`, the boids will stay together in a looser cluster, and at really low relative values of `c1` to `c2`, the boids can spread out over the entire domain or potentially for multiple barely distingushable clusters. To show the effect that `c1` has on the flock, I calculated the radius of gyration as a measure of size for the flock at different values of `c1` while keeping the rest of the scaling factors constant (`c2=0.01`, `c3=1`, and `c4=0.1`). Properly calculating the radius of gryation over periodic boundary conditions turned out to be more challanging than I expected!

You can see in figure 1, which shows the time averaged radius of gyration over the first 100 timesteps as a function of `c1`, that as the scaling factor increases, the radius of gyration decreases. Figure 1 also shows us that over the window that we are looking at (0.0000005 to 0.0000995), which almost spans 3 magnitudes, the relationship between time averaged radius of gyration is not quite linear, but instead the gain in compactness becomes less efficient as `c1` gets larger. Figure 2, shows the radius of gyration vs. time for many flocks of linearly increasing `c1`. The green lines represent flocks with lower `c1` and the blue lines represent flocks with higer `c1`. The fact that the blue lines are clustered closer together, despite the difference in `c1` from one line to the next being the same accross all the lines, shows that the relationship between `c1` and flock size is not linear.
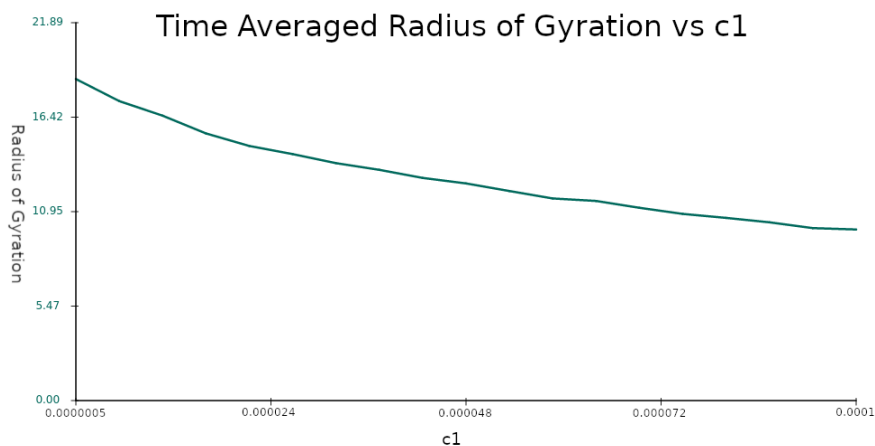

Figure 1: Plot of time averaged radius of gyration vs c1. The other scaling factors were `c2=0.01`, `c3=1`, and `c4=0.1`.
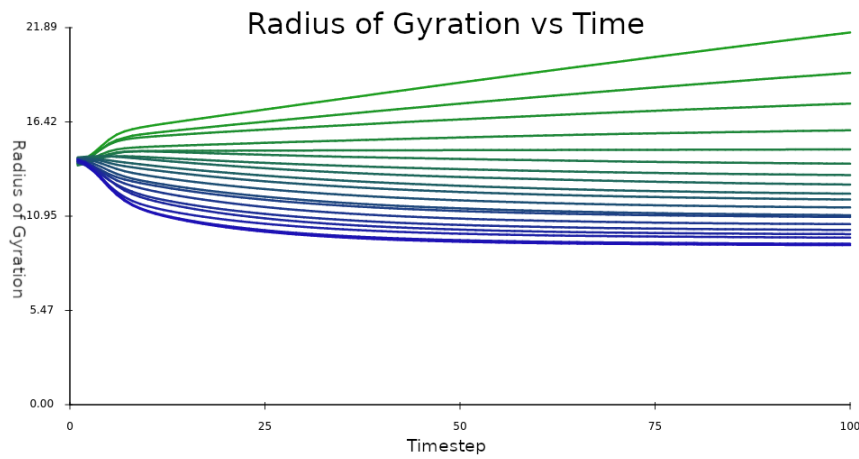
Figure 2: Plot of radius of gyration vs time for various values of c1. Green lines corespond to lower values of `c1`, and blue lines have higher values of `c1`. The other scaling factors were `c2=0.01`, `c3=1`, and `c4=0.1`.

c2

The second scaling factor is also known as the repulsion factor because it controls how strongly the boids will repel from each other. The effects of this scaling factor are virtually opposite of the effects of `c1`, so I won't bother explaining them with so many words, rather I will just include the same plots as I did for `c1` to show the difference. The range of `c2` covered is a similar almost 3 magnitudes (0.0005 to 0.0995), but is itself 3 magnitudes higher than the range of `c1` was due to the scale at which these forces inherently occur.

Figure 3 shows the time averaged radius of gyration as a function of `c2`, and pretty much shows the behavior we would expect, the relationship is much closer to linear than it was with `c1`, but a slight curve near the top hint's that it probably isnt. Figure 4 shows the radius of gyration as a function of time, where once again the lines of lower `c2` value are green, and the lines of higer `c1` are blue. The clumping of these lines also shows the almost-but-not-quite linear relationship between `c2` and flock size. Figure 4 also shows the behavior of the flock size rapidly expanding from the beginning until it reaches an equilibrium size, where it then slows down in expansion rate. This behavior is also reflected in Figure 5, which shows the time average alignment, a metric we will get to in the next section about `c3`, decreasing as `c2` increases. This is a result of the boids pointing away from each other for longer and longer times as they fly out to their large flock equilibrium size.
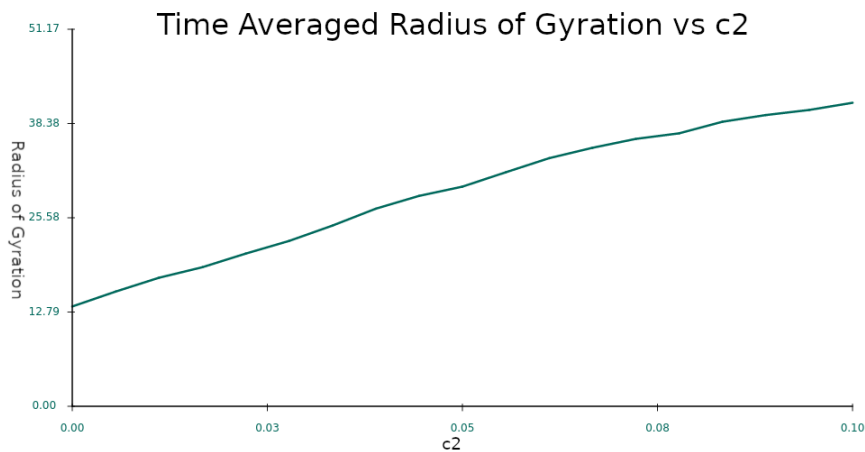

Figure 3: Plot of time averaged radius of gyration vs c2. The other scaling factors were `c1=0.00001`, `c3=1`, and `c4=0.1`.
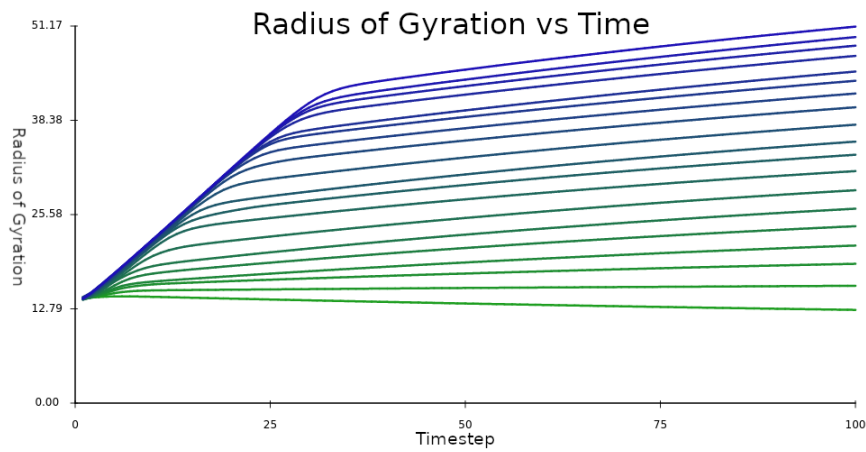
Figure 4: Plot of radius of gyration vs time for various c2. Green lines corespond to lower values of c_2, and blue lines have higher values of c_2. The other scaling factors were $c_1=0.00001$, $c_3=1$, and $c_4=0.1$.
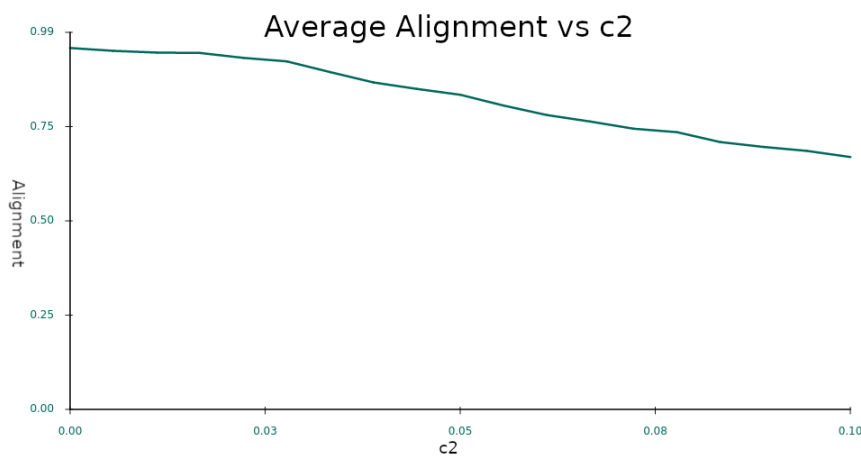

Figure 5: Plot of time averaged alignment vs c2. The other scaling factors were $c_1=0.00001$, $c_3=1$, and $c_4=0.1$.

The attraction factor, c_1, and the repulsion factor, c_2, can work together to cancel each other out. One can intuit that simultaneously increasing or decreasing the two by the same magnitude will yield a flock of similar size to before, but what does end up changing about the boid's behavior is their rigidity. I unfortunately don't have an official metric to measure this by, but at high values of both c_1 and c_2, the boids in the flock behave much more like molecules in a solid where they don't swap places with each other very easily. Whereas at low values of both c_1 and c_2, the boids in the flock look much more like they are molecules in a liquid or a gas, they can much more freely move around within the flock and aren't so constrained to their positions. This is partially because the relative strength of c_4 (the randomness factor) is a little higher when c_1 and c_2 are lower, but it's also because the ideal distance from boid to boid isn't so harsh and strictly defined when c_1 and c_2 aren't so high. To show visually this I have pictures of the flock after $timestep=100$ for two scenarios: 1. where $c_1=0.001$, $c_2=1$, $c_3=1$, and $c_4=0.1$, and 2. where $c_1=0.00000001$, $c_2=0.00001$, $c_3=1$, and $c_4=0.1$.
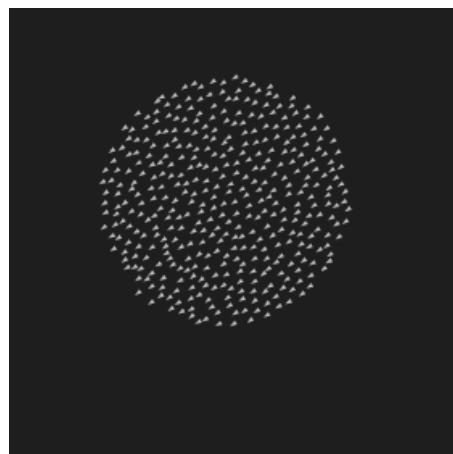

Figure 6: The flock forms a close to perfect circle when c_1 and c_2 are high, as the boids find their optimal positions quickly.

Figure 7: The flock flows around much more loosly when `c1` and `c2` are low, as the boids don't care so much about optimal positions.

c3

The third scaling factor is also known as the allignment factor because it controls how strongly the boids align themselves to the flock's average heading. Generally a lower `c3` value leads to more random wandering type movement from the flock whereas a higher `c3` causes the whole flock to align quicker and travel in a straigter direction. To demonstrate the effect of `c3`, I calculated the average aligment between each boid at varying values of `c3` while keeping the rest of the scaling factors constant (`c1=0.00001`, `c2=0.001`, and `c4=0.1`).

Figure 8 below shows the time average of this alignment measure plotted as a function of `c3`, and shows how order rapidly ensues out of the chaos that is low alignment factor, and then further increases in alignment gives rapidly diminishing returns as the flock quickly reaches near its maximum possible alignment of 1. Figure 9 shows the time averaged radius of gyration plotted as a funciton of `c3`, and it shows how there is generally just a trend towards a lower radius of gyration for higer values of `c3`. This makes sense as a more aligned flock will be able to compactify slightly better.
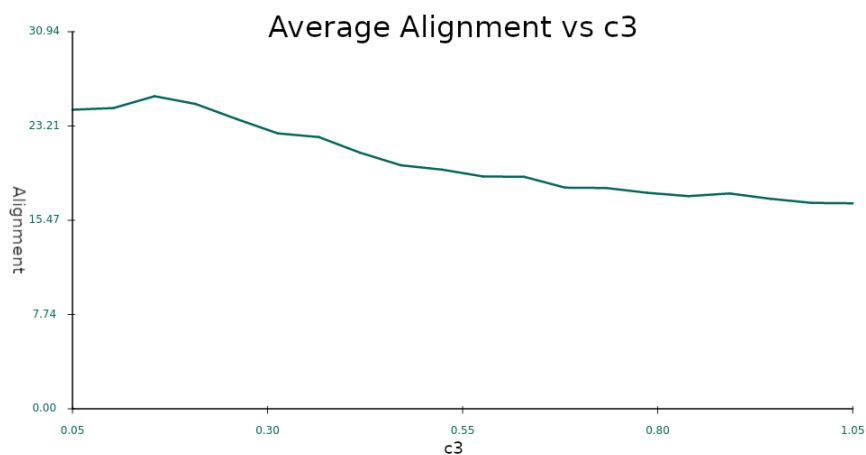

Figure 8: Plot of time averaged alignment vs c3. The other scaling factors were `c1=0.00001`, `c2=0.01`, and `c4=0.1`.
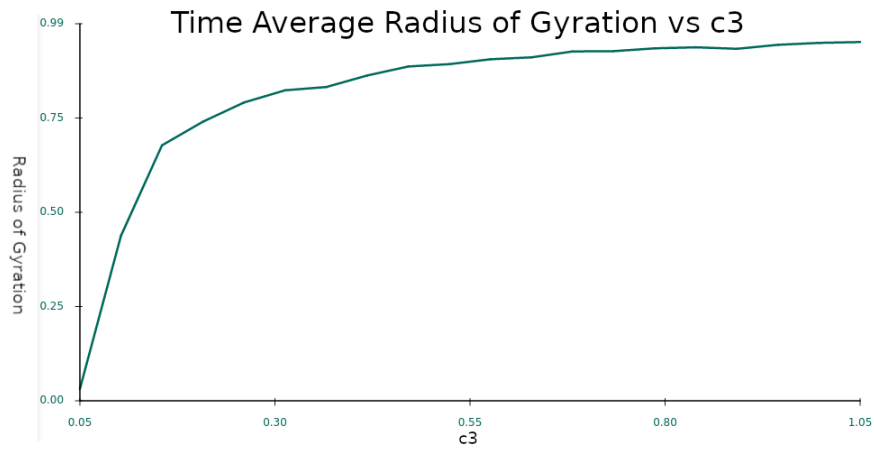
Figure 9: Plot of time averaged radius of gyration vs c3. The other scaling factors were $c1=0.00001$, $c2=0.01$, and $c4=0.1$.

c4

The fourth scaling factor is also known as the randomness factor because it controls how strongly random movement plays a role in the boids' movement. Similar to how $c1$ and $c2$ work as opposites in terms of flock size (as seen in radius of gyration), $c3$ and $c4$ work as opposites in terms of aligment. Figure 10 features a plot of the time averaged alignment, and shows how while $c4$ is smaller than $c3$ enough, the flock stays close to its maximum possible alignment. But as $c4$ approaches the same level as $c3$, the aligment of the flock drops, and shortly after $c4$ exceeds $c3$, the entire flock descends into chaos that can't really productively move anywhere.
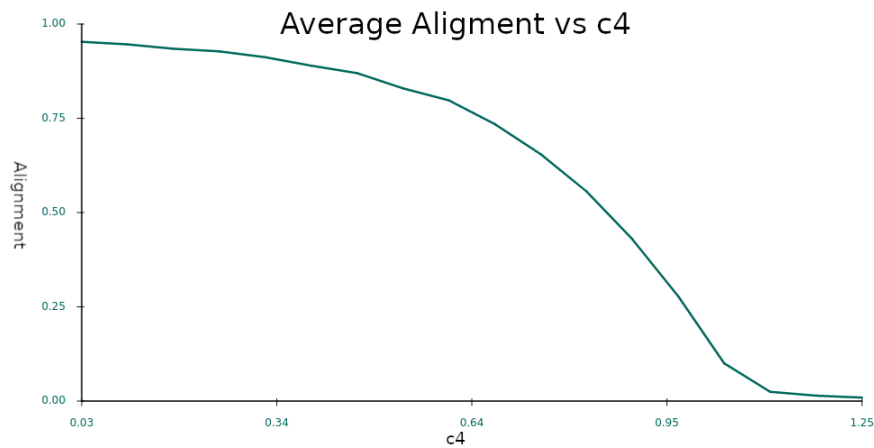

Figure 10: Plot of time averaged alignment vs c4. The other scaling factors were $c1=0.00001$, $c2=0.01$, and $c3=1$.

The fourth scaling factor can also have a slight effect on flock size, as a flock with more random movements will tend to spread out more, and take more time to reach its spread out state. This can be seen in figure 11, featuring a plot of radius of gyration vs time for the first 100 timesteps of the simulation, where runs with lower values of $c4$ are drawn in green, and runs with higer values of $c4$ are drawn in blue.
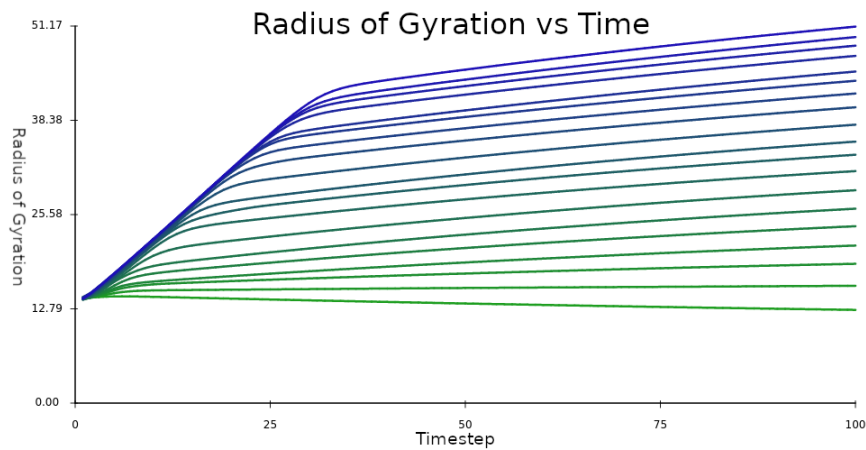
Figure 11: Plot of radius of gyration vs time for various c4. Green lines corespond to lower values of `c4`, and blue lines have higher values of `c4`. The other scaling factors were `c1=0.00001`, `c2=0.01`, and `c3=1`.

Obstacles

To introduce obstacles into the environment I created these 'repellant' objects that are created in random locations around the simulation arena. The boids see these repellant objects just like any other boid ecxecpt the only interaction they have with it is repelling, they skip using `c1`, `c3`, and `c4`. And the strength of the repel interaction is scaled by a "repellant strength" variable that I haven't come up with a good name for.

This repellant strength has a significant effect on how the boids interact with the repellants. With a relatively low number of repellants, 2-30, in a 200x200 simulation domain, and at a low repellant strength, around 1-25, the flock will generally try to avoid areas with more repellants, but usually the flock will fly over the repellants no problem and undisturbed. But at as the repellant strength starts to increase to about 30-200, the flock will avoid many more repellants, and when it does fly over one it will generally into pieces, maybe coming back together soon after. And as the repellant strength increases even further to really high values, 200+, the flock will just avoid flying into any repellants and just condense into a smaller and smaller flock size. At higher orders of magnitude, about 10^3, the flock starts to show signs of significant shrinkage due to there being such strong repellant forces in the environment. This is shown by the slight downward slope seen in figure 15 at 1000 and 10000. Still at really high values though the flock sometimes has no choice but to fly over a repellant, in which case it has a very high likelyhood of splitting because the repellant strength is so high. Eventually, at incredibly high repellant strengths (10^4+), the flock will break up into streams to find the most optimized spaces as far away from any repellants as they can get, and stay there. The significant upward slope in figure 15 at these magnitudes is caused by the flock's tendency to break up and find multiple local minima.

Figure 12: The flock doesn't really care so much about repellants when the repellant strength is low, in this case just 10. Boids are denoted with a triangle and repellants are denoted with a circle.
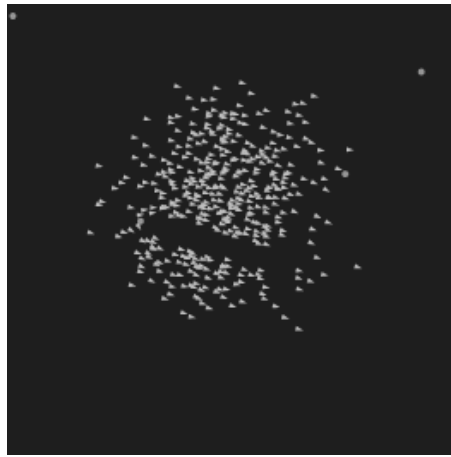
Figure 13: The flock will get lines "cut" through it by the repellants at higher repellant strengths. Boids are denoted with a triangle and repellants are denoted with a circle.
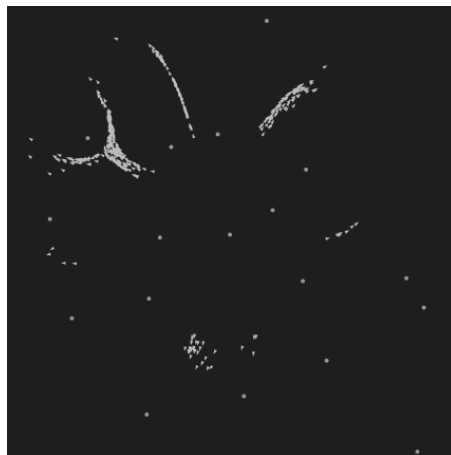

Figure 13: At very high repellant strengths (in this case 100,000), the flock will break appart into multiple streams to find local minima. Boids are denoted with a triangle and repellants are denoted with a circle.
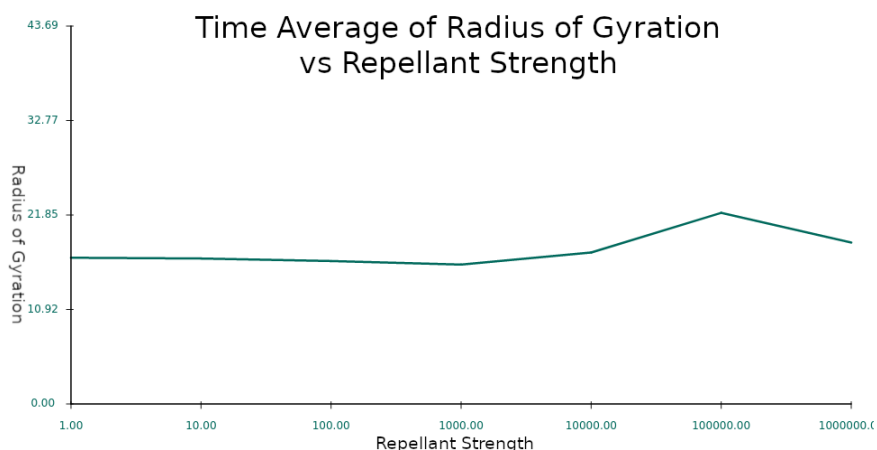

Figure 14: Plot of time averaged radius of gyration vs repellant strength. The scaling factors of the simulations were `c1=0.00001`, `c2=0.01`, `c3=1`, and `c4=0.1`.

Key Differences in the Model

My model is mostly super similar to the model presented in the starter code for this homework assignment, but there are a few differences. The biggest difference is that the boids update individually in my model as oppposed to all at once in the starter code. The reason behind this is just because that's how my object-oriented creative coding brain defaulted to, and so that's how I made it. It doesn't make a difference in the end result, maybe makes it slightly more lifelike, but it also means more computations are being done and so the code is less efficient. I did want to show off the fact that the boids update asynchronously though, so instead of (or maybe in addition to, the UI isn't complete yet) regular play/pause buttons, I created a slider that can speed up or slow down

the rate at which the simulation runs. At speeds less than 1, less than the whole flock is updated each frame, so you can see parts of the flock update at different times. I will be hosting the model as soon as I can at peter.rosenth.al/classes/csci-dynamic-models-in-biology/flocking. For now you can also view it by cloning the repository `https://github.com/peterrosenthal/csci-dynamic-models-in-biology`, and (assuming you have yarn installed on your system) build the HW3 project with `cd csci-dynamic-models-in-biology/HW3/`, then `yarn dev`, and `yarn serve`, then you can navigate to `localhost:8080` in your browser to view the project.

Paper Review: "Solving the shepherding problem: heuristics for herding autonomous, interacting agents" by Daniel Strombom et. al

Once again it was lovely to read about cute animals and their behaviours. I think that's one of the best parts about biology! The paper asserts that the algorithm(s) behind sheep - sheepdog interactions is relatively understudied, and that all previous models were quite far off. This is an incredibly important system however to have a deep understanding of, as it could have many real world applications, from crowd control to environmental cleanup. The major contribution that this paper is making to the field of shepherding is they are introducing a model that, according to them, is significantly less inadequate compared to all previous models.

The new model that the researcers introduce is both similar to the previous models and unique in its own way. All of the models, including this one, are flocking-based models, like we explored in our "simulation" section of this homework. They all model the sheep agents in a similar (greedy) way, where the presence of a shepherd agent drives individual sheep toward the center of the flock, and they all feature one shepherd agent that behaves differently from the sheep agents. The key principle that the researchers are exploiting here that makes their model so much better than the rest is that instead of expecting the shepherd agent to follow one single behaviour the entire time, they allow the shepherd agent to conditionally switch between modes where they exhibit different behaviors. The conditions that give rise to mode switching are how close the nearest sheep agent is to the shepherd agent, and if all of the sheep agents are within a radius of the global center of mass or not. The first mode, driving mode, is when the sheep agents are within that radius and the shepherd agent can position behind the flock relative to the target to drive the whole flock toward the target. And the second mode is is called collecting mode, when there are sheep agents that are too far away from the global center of other sheep agents, and the shepherd agent has to position behind that individual sheep agent relative to the global center of other sheep agents in order to drive it back into the flock.

The strongest part of the paper was the second figure that showed all the agents' traces of position over time. This was instantly and visually recognizable as resembling real shepherding to anyone that has even seen a single video of sheep - sheepdog interactions before. I love this plot a lot though, the inintial and final positions of the sheep agents show how first they were collected to a much more compact flock and then later after that they were driven towards their goal. And the trace of the single shepherd agent is clear and easy to see, and doesn't resemble anything too programatic or anything, it actually just looks *natural*. I think that this figure is so good and engaging that it should be featured as the first figure of the paper to excite the reader more, instead of the second figure.

The weakest part of the paper is that there were no pictures of smiling aussies, or even any pictures of any dog for that matter. I want to look at pictures of dogs at all times! and this paper is about dogs, so it should give me the chance to look at dogs! There are at least two different breeds of dogs mentioned througout the paper that I can find (border collie on pg.6, and Australian Kelpie on pg. 8), and while I sure know the difference between those two breeds, I'm sure there's plenty of other readers out there who would very much benefit from a picture of the dogs so that they too can know the difference.

The seccond weakest part to me, but a more serious weakest part, is that it is hard as a reader to actually compare how successful different models are in relation to each other. They start out by measuring the success of the base version of their model using the proportion of successful shepherding evets within 8000 time steps, but then later on when evaluating their "local shepherding" variant of the model (which they say is not as successful of a model yet), they introduce two more metrics to measure success by that I can't go back and compare against the base model myself to judge the difference. Furthermore they state that all previous models were much worse, and they qualitatively describe why they were worse, but the researchers never give any numbers or quantitative evidence to compare their model against previous models with, so it's impossible for the reader to actually know *how* much better the new model actually is. It would have been a much stronger paper I think if it had included the same evaluations, like proportion of successful shepherding events, of previous models so that there could be some real context to them describing how much better their new model is.

It was mentioned in the conclusions section of the paper that they only used one skilled sheepdog to collect their real world data from, and they would like to expand that to include a larger sample size of dogs, spanning different breeds and skill levels. I too would love to see the work expanded in that direction! I'm especially interested in how those two different parameter spaces, breed and skill level, turn out to affect the behaviour, and if they can be consistently incorporated into the model so that we actually can model more and less skilled dogs. Then perhaps we can experiment with tuning the skill parameter of the model passed what is normally observed in sheepdogs naturally, see what the "ideal sheepdog" could potentially be.