Find Stable Solutions to the Richardson Equation

CSE 589: Bioinspired Computing

Bryce Pedroza†  
 CIDSE  
Arizona State University  
 Tempe, Arizona USA  
bapedroz@asu.edu

Stephanie Forrest  
Biodesign Center for Biocomputation, Security and Society  
 Arizona State University  
 Tempe, Arizona USA  
steph@asu.edu

ABSTRACT

The purpose of this assignment is to find stable arms control scenarios using the Richardson Model. To this end, a genetic algorithm was written in Python 3 from scratch to find parameters that would lead to arms control among the three countries in the differential equations in the Richardson Model. This GA is represented using an object-oriented programming paradigm to simplify its design for end users. To determine its effectiveness, a combination of different metrics is observed. A novel normalization factor is proposed for to the fitness function in an attempt to promote more interesting individuals in the GA. This normalization feature is toggleable and its impact on results was also recorded. By including the normalization to the fitness function, more consistent results were reached that could produce the ideal arms control systems between the three countries. The result is an effective GA that can produce quality parameters that lead to fast arms control scenarios.

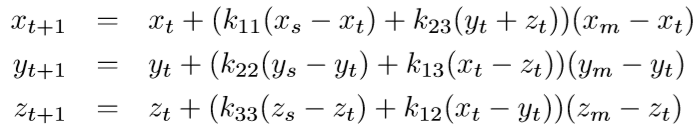
KEYWORDS

Genetic Algorithms, Richardson Equation, normalization, mutation, crossover, selection, Root-mean-square deviation, arms control

1. Introduction

Lewis Fry Richardson expressed an idea that the push and pull of arms spending was indicative of the outbreak of war. Therefore, he developed a system of differential equations that could estimate the probability of two countries engaging in war according to their arms expenditures and other fiscal constraints, thus known as the Richardson Arms Race Model. This model has been widely adopted to two country systems, but is there a reason it could not be adjusted to multiple countries? For the purpose of this assignment, the Richardson Equation was updated to reflect the introduction of a third country that can contribute to the outbreak of war. This iteration more accurately reflects the advanced international economy of the world today (USA, Russia, and China, for example). In this version, the two weaker countries become allies, and pool their resources against the dominant country.

This model can exhibit some strange behaviors. First, there is the possibility of a shift in alliance. This would be the result of one of the weaker countries overtaking the dominant country, thus causing a new alliance between the new set of weaker countries. Secondly, instability within the nations can present itself in the form of a runaway arms race in which the countries continue to increase their spending on arms over each time step. Ideally, the countries would find a scenario in which the spending remains constant and shifts in alliance no longer occur. This would be considered as arms control and is the steady state proposed in the Richardson Equation.

****

**Eq 1.** Richardson model expressed in three differential equations

However, there are 12 parameters that can be adjusted in this equation, so it can be difficult to manually determine how these values can be set such that a state of arms control between the three countries can be achieved. To do this, a genetic algorithm can be created to produce the possible parameterizations.

2. Methodology

The genetic algorithm was implemented Python 3 without the use of genetic algorithm libraries. The parameters for this GA are all real numbers between 0 and 1. The GA is designed to create a variable size population with random initial parameters, perform crossover between individuals, mutate individuals, and perform a selection tournament to pick the fittest individuals. By performing crossover and mutations during each iteration of the genetic algorithm, the search space is continuously explored, and genetic diversity is increased, thus leading to more fit generations over time.  These actions are performed over multiple generations with the goal of increasing the fitness of the overall population over time.

To perform mutation, one value for the individual is randomly chosen. This value is then increased or decreased by 5% of its current value. Now, these are real values and must be in between 0 and 1, so if the mutation was to increase or decrease the value above or below these thresholds, mutation would be limited, and the value would be kept within this predefined range. There is a variable mutation rate that can limit the number of mutations in a generation.

Crossover switches one value from two randomly selected individuals in the population. There is a variable crossover rate to limit the number of crossovers performed in a generation. The crossover is done with replacement. Thus, an individual can be crossed over more than once.

The selection tournament is a one versus one selection tournament with replacement, thus allowing for an individual to be added to the next population more than once. The fitter of the two individuals is selected. The fitness of these individuals is calculated using a fitness function. An individual is more fit if the current arms spending between the dominant country and the two allied countries is close to 0.

**Eq 2.** Standard fitness function for Richardson Model

3. Representation

The values for the GA are being represented using Python objects. The lowest level object is a Country(). Each Country holds each of their four unique parameters that are randomly assigned, mutated, and crossed over. Along with this, each country holds its current arms spending and whether or not it is the current dominant country. A country has the ability to take in all of the above parameters for the purpose of debugging possible parameterization.

The Richardson() object simulates a single individual within the GA population. Each Richardson contains three countries, x, y, and z. Along with these three countries, it contains its fitness and mutation rate. The Richardson class can take all of these as parameters, again for debugging purpose, but by default it uses a static method to randomly create the three countries and assigns its initial fitness to -1. The mutation rate must be passed into each Richardson object. Its publicly available functions are to mutate and to perform fitness calculations. Mutation is performed as described in the previous section. To perform fitness calculations, Eq. 1 is performed dynamically depending on which country is currently dominant. After the spending is calculated, the countries are reassessed, and a new dominant country is assigned if one of the allied countries has started to outspend the dominant country. Once any possible reassignments are complete, Eq. 2 is performed and assigns the fitness for the individual.

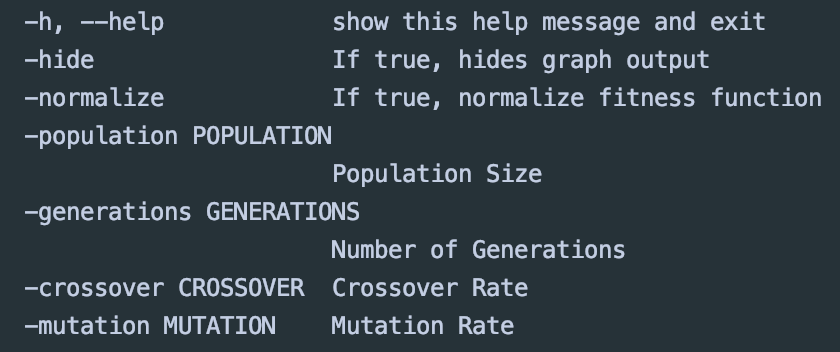
The GeneticAlgorithm() object contains all the information needed to perform the entire GA as the name suggests. There are four required parameters for the GA: population size, crossover rate, mutation rate, and a normalization flag (this flag will be addressed later). Upon initialization, the GA creates a population of Richardson objects according to its parameterization. It has one publicly available function, which is to iterate the population. This function performs crossover, mutation, selection tournament, and historical tracking, in this order. Mutation and crossover have been described in detail in previous sections.

However, there is a small caveat when performing the selection tournament. When adding the fitter individual to the next generation, a deep copy for that individual is sent and not the individual itself. A deep copy will create a new object with the same attributes as the original. If a deep copy is not used, then the new generation will have multiple references to the same individual in the population. This means, without the deep copy, if one individual is crossed over then that change is reflected everywhere that object is referenced in the population. This could cause individuals to be mutated or crossed over significantly more than intended.

Once the selection tournament is performed, the GA object performs some additional calculations to track the best global solutions and the average fitness of the population for each generation. With these objects in place, the GA can be run quite easily for a variable number of generations to observe how performance changes.

4. Parameterization

Because the final GeneticAlgorithm() object takes in 4 parameters and there is one additional parameter for the number of generations that will be performed, there are the 5 values that can be adjusted from run to run. These values can be adjusted with command line arguments to allow for quick parameter adjustments without having to touch the source code. Below are the parameters that can be adjusted at the command line. A full detailed list of ways to run the code can be found in the README.md

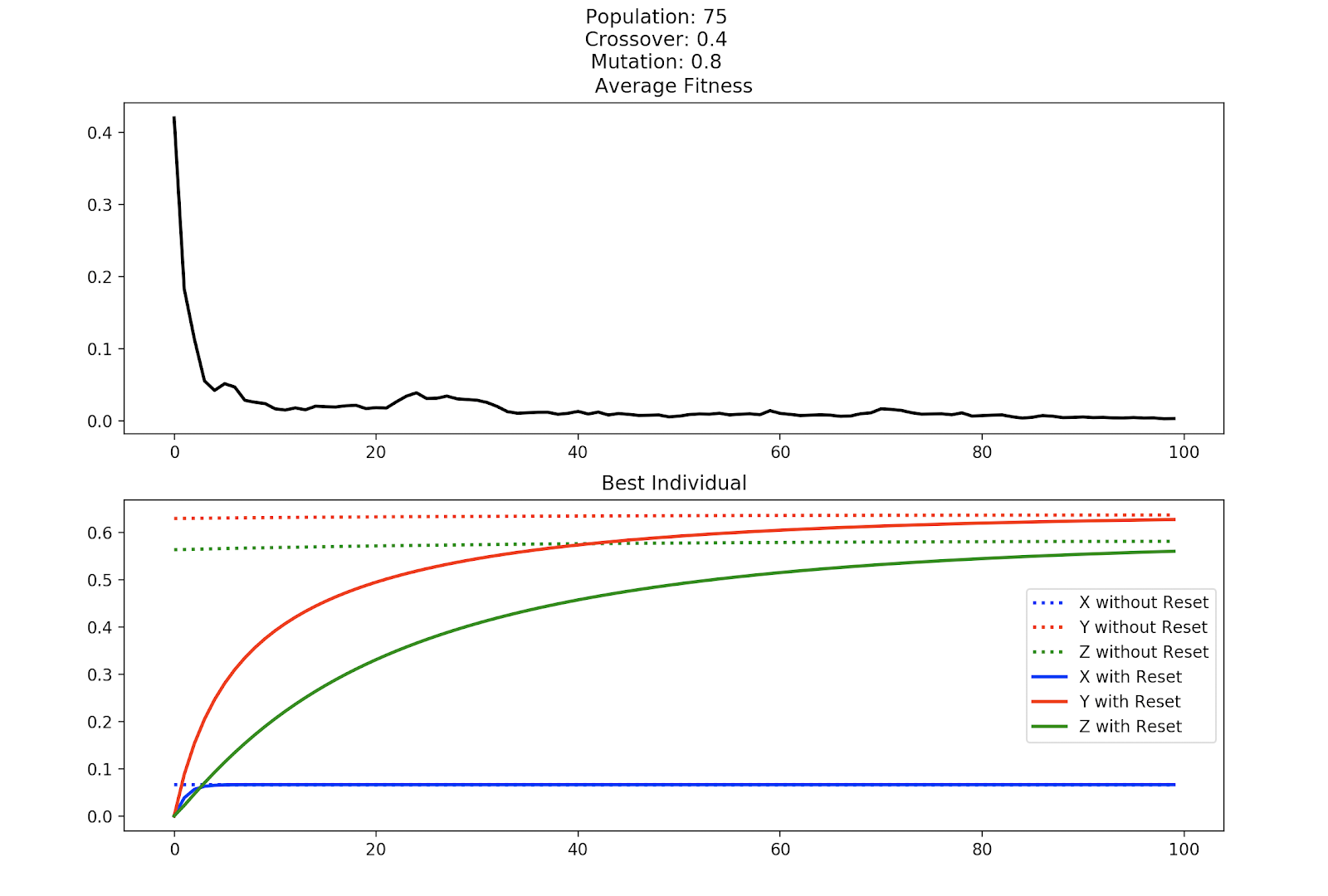


**Fig 1.** Here are the possible command line arguments that can be used when running the GA. These params will adjust the main parameters for the GA. The appendix will include additional details on how to run the code.

5. Experimental Design

Once a GA runs for the specified amount of generations, a way to assess the results had to be produced. A visualization was created to observe any patterns that might have manifested in the final results. The first was a fitness curve to show average fitness across each generation in the GA.

The next visualization was built to study how the best individual’s parameters performed over time.  The individual that produced the highest fitness was taken from the population and simulated again with its countries’ parameters and where *t* is the number of generations simulated. These same parameters were used again but initial spending was set to 0, . This was done to compare how the parameters would perform according to the initial state assumed in the assignment. These calculations were recorded for the number of generations ran by the GA, giving the Richardson Equation enough time to accurately observe its expected results vs its actual results.



**Fig 2.** Here is a standard output from the GA. The first visualization shows how the average fitness for the GA converges towards 0, which is ideal. There are some perturbations in the average fitness that can be attributed to the GA mutating and performing crossover to promote genetic diversity and to increase the search space. The second visualization shows the best individual and its counties’ spending over time. The dotted lines assume initial conditions . The solid lines assume initial conditions . The faster the actual results can simulate arms control, the better the final solution is.

Root-mean-square deviation (RMSD or RMSE) is used to measure how the expected model behaves () vs. how the actual conditions behave (). RMSD can thus be used to determine how the fitness and country arms spending of the actual function compare to that of the expected function. A low RMSD indicates the differences between the actual and expected values are small. RMSD can be calculated for fitness as well as for the countries’ average spending. RMSD for fitness is a measure of how consistent the fitness is over time. RMDS for the countries is a measure of how fast the countries converge to their expected arms control point.

However, when running ad hoc experiments to observe how parameters could impact performance, the best individuals were typically spending an extremely limited amount on arms. This would result in them having incredibly low fitness, but their RMSD measures would be disproportionately higher, meaning their parameters did not actually perform well. These individuals rise to the top because of how the fitness function promotes extremely low spending on arms. Consider . Using Eq. 2, this fitness, F, is 0.01. Now consider . This tuple would produce , a fitness 5 times worse than the previous. However, if the two fitness were normalized, the latter is better than former by nearly a factor of two. So, while the second could have relatively a more stable solution, the current fitness function does not promote that behavior.

To address this problem faced in the fitness function, the ability to normalize the fitness function was proposed. (Adjusting the fitness function might have been out of scope for this project but normalizing the fitness function did provide much better results) By standardizing the fitness function, low scale results are evaluated fairly against the larger spending results. When enabled with the normalization flag mentioned in section 3, the fitness function is updated to Eq. 3, eliminating the problem addressed in the paragraph above. With Eq. 3, would produce and would produce , accurately reflecting that the latter is closer to arms control. To further support normalization of the results, the normalized root-mean-square deviation (NRMSD) was utilized to evaluate the results.

**Eq 3.** Normalized fitness function

**Eq 4.** NRMSD, where y is either fitness or country spending

Experiments were designed to study how population size, number of generations, mutation rate, crossover rate, would impact the average fitness of the GA, the best individual fitness produced by the GA, the country NRMSD, and the fitness NRMSD. Additionally, with the introduction of the normalization technique, running trials with it enabled and disabled should help in understanding how it impacts the overall results from the program. To assess all of these different possible combinations, a standard baseline of values was set. The values in the baseline appeared to give decent results before beginning official experiments. Thus, adjusting from these baseline values should yield results that are acceptable but still show how parameters can affect different parts of the output. Each set of parameters is run 10 times, and the average across all the runs is recorded. These 10 runs are classified as a single experiment.

To evaluate how consistent the results are in a single experiment, the standard deviation was also recorded for the country NRMSD and fitness NRMSD across all runs. This should assist in understanding how the results vary between each run in the experiment.

6. Experimental Results

The results can be found as figure 1 in the appendix, and the entire recorded results for all the runs can be found in the excel sheet in the submitted zip file. Here is what was observed. The parameters that produced the results will be displayed in order of [**Population, Crossover, Mutation, Generations, Normalization]**. The baseline was determined to be ***[100, 0.8., 0.4, 100, True]***

For average fitness, the best results were seen with *[75 0.4 0.8 100 False]*. However, it should be noted that with normalization set to false, Eq. 2 was utilized for the fitness function. Thus, individuals with lower spending would be classified as fitter than those spending more even though the lower spending results might not lead to arms control. The worst fitness was with *[50 0.4 0.8 100 True].* This makes sense because smaller population can lead to less coverage in the search space. This can mean it takes longer to find the best fit. This same run, *[50 0.4 0.8 100 True]*, also had the worst best fitness overall, which again makes sense due to the small population size. *[100 0.3 0.8 100 False]* was found to have the best overall fitness, which again makes sense as again normalization was disabled.

*[75 0.4 0.8 100 True]* found to have the best country NRMSD and fitness NRMSD across all different experiments. In fact, all experiments in which normalization was enabled had a lower country NRMSD and fitness NRMSD than all runs where normalization was disabled. Therefore, the best individuals for the normalized runs consistently converged to the expected values while the non-normalized experiments were significantly higher, thus suggesting a low correlation between the expected arms control and the actual results. Figure 2 in the appendix will show what a run with small arms spending actually performs against its expected.

But how much did these runs vary? In terms of the most consistent result for country NRMSD, *[100 0.4 0.8 100 True]* produced a standard deviation of only 0.1093963639.  *[100 0.3 0.8 100 True]* gave the best results for fitness NRMSD, with a standard deviation of 0.106332323. This appears to suggest that mutation rate is more important than crossover rate in impacting how the results from each experiment change because even though the crossover rate was changed by a factor of 0.1, the best standard deviations did not vary too much. This is further supported by the runs that had the worst standard deviations. Country NRMSD *[100 0.45 0.75 100 True]* had a standard deviation of 0.3439660882 across its experiments, nearly three times worse. Fitness NRMSD had a similar result of 0.307012628 for *[100 0.4 0.7 100 False]*. These can show that decreasing the mutation rate will cause a decrease in the search space, causing runs to converge to different local optimums.

The results do not indicate that there is an end all set of parameters that perform better than all other options, but the results do indicate that parameters can be tuned to find the results that fit different use cases. For overall fitness, removing normalization can prove effective. Including normalization will increase the relationship between the expected results of the parameters and the actual results. Keeping the mutation rate high will ensure a higher search space, causing a decrease in the variation of results from run to run. If speed is being considered, decreasing the number of individuals in the population could prove to be successful while not impacting results significantly. However, it would be best to keep the population above 50 individuals to avoid poor results.

7. Conclusion

The GA implemented was able to produce consistently fitter individuals over time. The quality of the individual did vary depending on whether normalization was considered in the fitness equation. Results that did have normalization more consistently converged to the arms control observed in the expected results, while those without normalization often had the desired low fitness score but did not have the consistent low fitness assuming initial conditions and instead found more converging arms control than expected. Adjusting mutation rate and crossover rate impacted how the results varied across different experiments, and as population sizes decreased, the average fitness across all individuals should be expected to fall. The difference in results by varying the number of generations in an acceptable range (50-100 generations) did not have a large impact on any of the results, but lower number of generations will not give enough time to produce meaningful results with the GA.

REFERENCES

[1] Patricia S. Abril and Robert Plant, 2007. The patent holder's dilemma: Buy, sell, or troll? *Commun. ACM* 50, 1 (Jan, 2007), 36-44. DOI: <https://doi.org/>10.1145/1188913.1188915.

[2] Sten Andler. 1979. Predicate path expressions. In *Proceedings of the 6th. ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages (POPL '79)*. ACM Press, New York, NY, 226-236. DOI:https://doi.org/10.1145/567752.567774

[3] Ian Editor (Ed.). 2007. *The title of book one* (1st. ed.). The name of the series one, Vol. 9. University of Chicago Press, Chicago. DOI:https://doi.org/10.1007/3-540-09237-4.

[4] David Kosiur. 2001. *Understanding Policy-Based Networking* (2nd. ed.). Wiley, New York, NY..

APPENDIX

Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here. Insert paragraph text here.