Exploring Kin Directed Altruism in Populations of Evolving Neural Networks

University of Sussex
Artificial Life
Joshua Dare-Cullen

Abstract

This paper explores the idea of altruism, where by an individual decreases its own fitness at the increase of another individual. The idea of kin direct altruism is encoded into the input of a 3 layer artificial neural network in the fight for reproduction. A more natural approach in evolution is taken by having agamic reproduction of sisters happen at the end of each generation. The concept of genetic relatedness is explored to analyse the boundaries of genetic similarity is needed in the parameters of the neural network of related sisters, in which the expected egotistical behaviour of not giving food is dominant in the population.

1. Introduction

The idea of altruism has been seen by many, including Darwin himself, as the antithesis of the theory of natural selection, where individuals seemingly reduce their own fitness to increase others [1]. The emergence of these characteristics forced Darwin to the study of Beehives, where he found sterile worker bees helping their blood relatives, especially the queen, by giving blood. Darwin came to the conclusion that altruism may favour selection within related groups [2]. Since then, this apparent paradoxical behaviour has been discovered throughout other species and has been documented in over 220 bird and 120 mammal species, where acts of altruism are displayed through self-sacrificing individuals forsaking reproduction in pursuit of aiding siblings and/or parents [3]. Although, this level of altruism is considered low in comparison to the extreme altruism cases of fatal self sacrifice in Eusocial insect colonies [4].

The theory of Kin-selection developed by *Hamilton* in 1964 [5] and expounded upon by *Trivers* in 1985 [6], has been introduced to explain why, contrary to *survival of the fittest*, genetically-based altruism is observed in many organisms. Hamilton derived that altruistic behaviour will be maintained in a population if it is directed towards one's own kin. The fitness of the altruistic individual will be decreased by the altruist act but the fitness of the kin individual who is the recipient of the act will be increased [5].

In more recent studies within the theory of psychological egoism it has been suggested that no act of sharing, helping or sacrificing can be described as truly altruistic, as the entity may receive an intrinsic reward in the form of personal gratification. Although, the validity of this argument depends on whether intrinsic rewards qualify as "benefits" [7]. Throughout this paper we refer to the artificial simulation research work by *Parisi et al*. [8]. Specifically, the experiments conducted on kin-directed altruism in social simulations. Where, it was demonstrated with a population of 100 minimalist 3 layer artificial neural networks which represented 20 groups of 5 sisters, the related sisters would act altruistically towards each other and egotistically towards non related sisters with the incentive to reproduce at the end of an evolution cycle. The measure of altruism was based on food and governed by the neural network. Each sister would have a fixed amount of social encounters between kin and non kin in which it would have to decide whether to gift (0) or not gift (1). To use Hamilton's [5] theory of genetic relatedness as the driver of altruistic behaviour, the sisters were genetically related by the weights that made up the neural network. The neural network architecture remains the same but what differs is the introduction of biases and non linearity in the hidden layer and output layer. This enabled extra complexity to the decision making of the sister but also to stretch the genetic relatedness the sisters have to maintain to exude kin directed altruism. Unlike the random amount of weights selected for mutation at every evolution cycle as in *Parisi's* simulation, all weights and biases of the sisters were mutated in the simulations within this paper. No specific mutation rate was specified in the original experiment and a single population size of 100 was used. Regarding this, a varying of mutation rates and population sizes were tested in order analysis the threshold of genetic relatedness needed. What was discovered that if the weights are mutated above a specific threshold, the individuals are not related enough in the parameters of the neural network to evolve altruistic behaviour. Although, kin directed altruism was analysed in simulations in which kin where the only receivers of food and no non related individuals ever received food from each other. So, a split between altruistic and egotistical behaviour became 50/50.

2. Methodology

2.1. Neural Network

The sister's decision to act altruistic or egotistic is governed by a feed forward neural network (Fig. 1). It has the architecture of 2 input nodes, 3 hidden nodes, and 1 output node. The input into the network is a vector with two possible binary combinations. For a social encounter with a kin, that is the genetic relative, the input is encoded as (1, 0), e.g. ($x1$,$x2$) in Fig. 1. The input for a social encounter with a non kin, a genetically different network, is encoded as (0, 1).

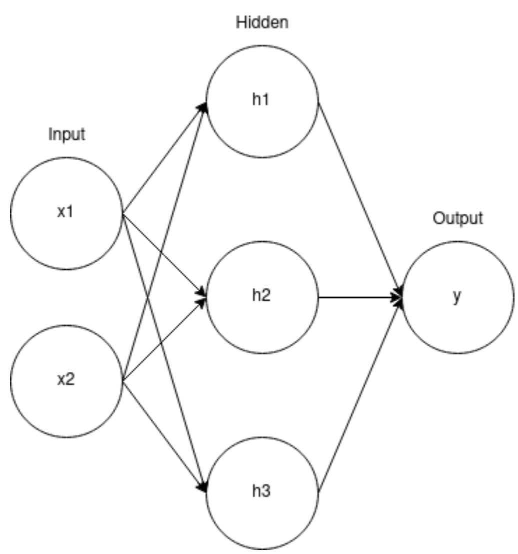

Figure 1. The neural network architecture, with the input flowing from left to right.

The hidden layer comprised of a Sigmoid activation function, discovered by Rumelhart [9] and seen in Fig. 2, it scales any real number input between 1 and 0. This introduces non-linearity into the network where 1 represents the neuron being fully activated and 0 not activated.

$$sigmoid = \frac{1}{(1+exp^{-(h)})}$$

Figure 2. The sigmoid logistic function, used in both the hidden layer and the output layer as the activation function.

$$h_1 = sigmoid(w_1^T \cdot x + b_1)$$
$$h_2 = sigmoid(w_2^T \cdot x + b_2)$$
$$h_3 = sigmoid(w_3^T \cdot x + b_3)$$

Figure 3. Hidden layer equations, x is the input vector encoding the social encounter.

The equations in Fig. 3 can be reduced down to a single equation in Fig. 5. Where W is a 2 by 3 matrix composed of the weights that are used to execute the dot product of the input, with the matrix requiring a transpose. b is a vector with a length of 3 and is added to the 3 resulting calculations of the dot product. The dot product of W the weight matrix and the input vector x can be mathematically represented in Fig. 4.

$$W^T \cdot x = \sum_{i=1}^{N} w_i x_i$$

Figure 4. The dot product expressed as a matrix product.

$$h = sigmoid(W^1 \cdot x + b^1)$$

Figure 5. The succinct representation of the hidden layer in the neural network.

$$y = sigmoid(W^2 \cdot h + b^2)$$

Figure 6. The output layer of the neural network. The matrix W^2 takes the shape (1,3) and the bias b^2 is a single value.

Upon the calculation of the dot product with the output of the hidden layer nodes and output node will be encoded between 0 and 1 with a probability of it either acting egotistically (0) or altruistically (1). The threshold boundary for the value being rounded to 0 or 1 is 0.5, that is if the output is greater than 0.5 then the output is considered altruistic and is rounded to 1. If below 0.5 then the value is rounded to 0.

2.2. Evolution Algorithm

Every individual within the population had the weights and biases of its network randomly initialised to values between 0-1 from a Gaussian distribution and represented by a 64 bit floating point number. Each group of 5 sisters in the 20 possessed the same randomly generated weights and biases, a reflection of their genetic relatedness. There were 50 generations and every generation contained 100 encounters for each individual in the population. Every individual started with the same amount of food which was equivalent to the amount of social interactions, 100 pieces of food to either give or keep in encounters with either kin or not kin. If the individual decides to give a piece of food away then 1 is decremented from its food total and 1 given to the recipient individual in the social encounter. The probability of a sister meeting a kin is setting by random shuffling the population and drawing two individuals. At the end of the 100 encounters for every individual the top 20 are selected for reproduction based on the number of food in their possession. This type of selection is taken directly from *Parisiet al*. [8] and is the fitness function. The reproduction is agamic, hence an individual reproduces one or many offspring. The initial steps for reproduction were taken from *Parisiet al*., where the top 20 individuals of each generation would create 5 copies of itself with random mutations being applied to a random selection of weights. Making only the first generation of sisters having identical parameters in the network. The result of making 5 new bloodlines for each generation was that a number of simulations resulted in individuals giving to both kin and non kin at every encounter. This was because populations that acted more altruistically obtained higher amount of individuals in the top 20. For example, if 4 groups of sisters all make it into the top 20, the next generation the population would contain a fifth of each of the 4 genetically related groups from the previous generation. The problem is with the next generation, the 20 are made up of 4 separate blood lines. So, will give altruistically to an individual even though it is not technically related in the simulation but are similar in network parameters. The method used instead was for the offspring to inherit the network parameters of the parent as well as the blood line. This meant that a population had the incentive to get more individuals in the top 20. Also, this method allowed for tracking the networks from the the 1st generation to the 50th to see which ones prevailed. The probability of encountering a kin grew in proportion to the amount of individuals of the same bloodline were in the total population. Although similar to that a steady state genetic algorithm, the process of reproduction and mutation is derived from *Cecconiet al*. [10]. There is no cross over and one individual can produce many offspring. As mentioned in Sec. 1, mutation was applied to all of the weights and biases within the network of the individual. In our simulations a random number was generated to either take or add to each weight and bias from a Gaussian distribution with a set upper threshold. This threshold was set and varied by the researcher. The higher the threshold the less genetically related the offspring could potentially be. To gather a more accurate representation of the trajectory of genetically related individuals, the simulation of 100 social encounters with 50 generations was run using a varying set of mutations rates and population sizes as shown in Fig. 7. All combinations of population size and mutation rate were run for 10 iterations and 5 different random seeds. The increase in population did not effect the proportion of the 20 groups of sisters in the population. Only the amount of sisters in each genetically related group. The cut off for an individual to be reproduced was fixed at 20 for all the simulations conducted.

Mutation rates: $[0.0001, 0.001, 0.01, 0.1, 0.2, 0.3]$
Population sizes: $[100, 200, 300, 400, 500]$

Figure 7. The varying parameters, each population and mutation rate combination was tested over 5 different seeds.

3. Results

3.1. Measuring Genetic Relatedness on Altruistic behaviour

The altruism towards kin can be seen in Fig. 8, showcasing when the mutation value threshold is higher the less altruistic individuals become to their sisters. In both simulation this can be observed. As seen in Fig. 9, beyond 0.2 all individuals become egotistical in every social interaction, not giving up any of the food and defaulting to a fitness score of 100. The reward of giving to kin is not there because the network parameters are not similar enough. Every individual is acting autonomously and selfishly in pursuit of reproduction.

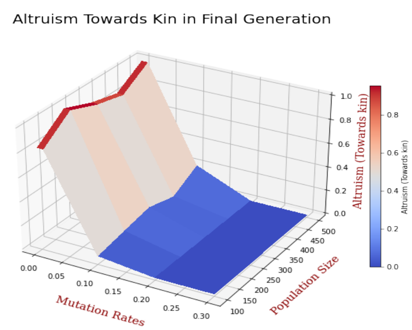

Figure 8. Mean altruism towards kin of the top 20 fittest individuals computed using results from 10 runs on all combinations of mutation rate and population sizes.

In regards to the increase in kin directed altruism, all population sizes and mutation rates from 0.0001 to 0.1 see a dramatic increase is in altruistic behaviour towards kin, as the mutation rate approaches 0 the individuals give a piece of food at every social encounter with kin and giving very few amounts of food to non kin. At a population of 200 and a mutation rate of 0.0001, it can be observed that the individuals in the final generation did not act altruistically at all to non kin, but instead, acted altruistically in every encounter with a sister. What can be analysed in both Fig. 8 and Fig. 9 is that there is more individuals in the population acting altruistically to sisters who have a similar set of

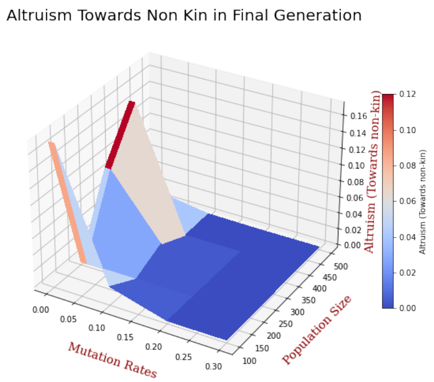

Figure 9. Mean altruism towards non kin of the top 20 fittest sisters computed using results from 10 runs on all combinations of mutation rate and population sizes .

network parameters than non sisters who have a different set of network parameters. From deriving the most altruistic individuals in Fig. 9 and Fig. 8 taking the parameters: *mutation rate* = 0.0001 and *Population* = 200. At around 15 generations the individuals change from being altruistic towards non kin, to kin. Every generation the altruistic behaviour improves towards kin except the period between 20 and 25 generations. Then it continues in a linear trajectory and plateaus at the max amount that can be given in one generation. The max fitness is achieved at 30 because one type of genetically similar weights (blood line) has dominated the whole population. This is shown by calculating the number of blood lines in each generation of 50 that can be seen in Fig. 11.

It can be observed that the amount of varying blood lines in the population decreases rapidly. The first 5 generations of evolution more than halves the variety of neural network parameters are left in the population. Interestingly, the networks that act altruistically towards each other in the initial phases of evolution is the bloodlines that go onto dominate. Between the generations of 10 and 20 the remaining 5 blood lines are fighting for the chance for reproduction. It eventually settles on 1 bloodline in the population.

The parameter for the amount of individuals proposed by *Parisi et al*. [8] was 100. The *mutation rate* is 0.0001. The reconstruction can been seen in Fig. 12. The trajectories are similar to *Parisi et al*. *'s* but the cross over of trajectories in ours occurs 5 generations after, at 20. This can be put down to the need for more generations to evolve altruism because the genetically related sisters have every value different in

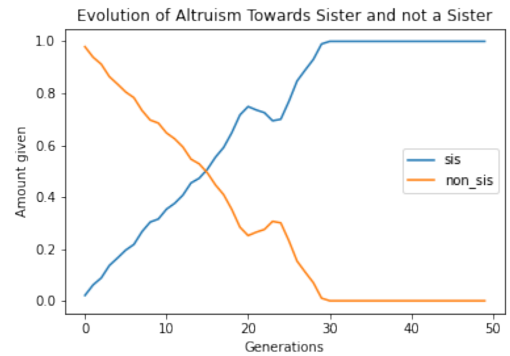

Figure 10. The mean number of altruistic and egotistic acts in the population, taken from 10 runs and 5 different random seed environments. The amount given is scaled by the maximum that it can be 100.

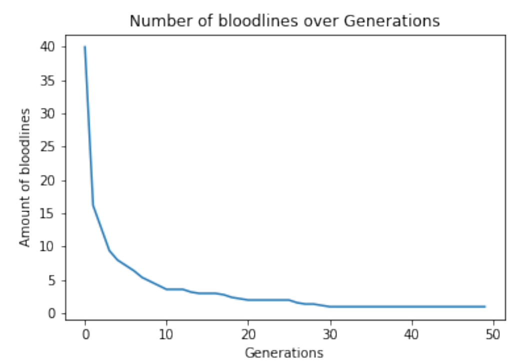

Figure 11. The mean amount of groups of sisters (bloodlines) within the population over 50 generations.

their networks, not one network in the population has the same set of network parameters. Fig. 13 shows that not one sister bloodline (group) dominates as in Fig. 11, two does.

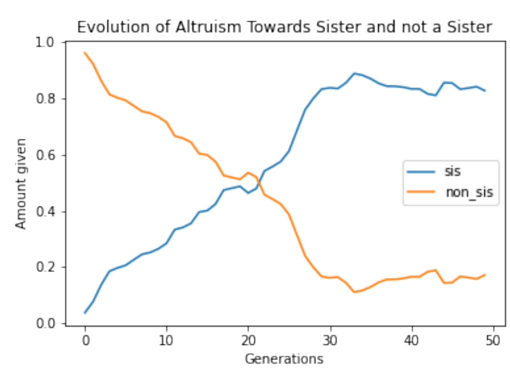

Figure 12. Mean.

The number of bloodlines when the population is 100 shows a rapid decrease from generations 0 to 5, similar to that of the graph in Fig. 11. What is differing is the generation from 5 to 30, a much slower decline and more bloodlines remain in the population deeper into generations.

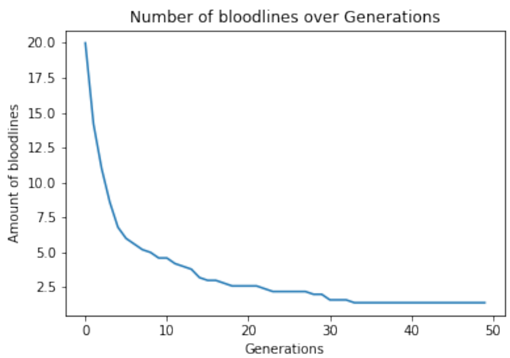

Figure 13. The amount

4. Discussion

The simple artificial neural networks simulated in this paper have shown the ability to evolve kin direct altruistic behaviour. Although, the conditions for this to occur is a small discrepancies in the neural network parameter values of the related groups of sisters. What this simulation doesn't highlight is the spacial nature of humans in prehistoric times and how it has played a pivotal role in shaping the evolution of altruistic behaviour. In contemporary relationships are not fully reliant on the spacial vicinity humans that makes one act altruistic. The influx of telecommunications has created a whole new way are interacting and maintaining relationships. The idea of cultural transmission is much more harder to replicate in artificial social simulations where imitation is drawn from cultural experience. The need for research is needed to truly understand the nature of altruism in natural world.

References

[1] A. Domondon. A history of altruism focusing on darwin. *Endeavour*, 37(2):94–103, 2014. 1

[2] C.R Darwin. *On the origins of species: By means of natural selection or the preservation of favoured races in the struggle for life*. Oxford Press, 1859. 1

[3] L Lehmann T. Kay and L. Keller. Kin selection and altruism. *Current Biology*, 29(11):438–442, 2018. 1

[4] F. Ratnieks and T. Wenseleers. Altruism in insect societies and beyond: voluntary or enforced? *Trends in Ecology and Evolution*, 23(1):45–52, 2008. 1

[5] W.D Hamilton. The evolution of social behaviour. *Journal of Theoretical Biology*, 7(1):1–16, 1964. 1

[6] R. Trivers. *Social evolution*. Benjamin Cummings, 1985. 1

[7] N. Burton. *Heaven and Hell: The psychology of the emotions*. Acheron Press, 2020. 1

[8] N. Gilbert, R. Conte, D. Parisi, F. Cecconi, and A.Cerini. Kin-directed altruism and attachment behaviour in an evolving population of neural networks. pages 238–251, 1995. 1, 3, 4

[9] David E. Rumelhart and James L. McClelland. *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: foundations*. The MIT Press, 1986. 2

[10] F. Cecconi and Parisi. Neural networks that learn to predict where food is and also to eat. *National Research Councils*, 1989. 3