The Evolution of Complex Attributes in a Species of Simulated Agents

Jay B. Nash, Gary B. Parker, Jim O'Connor
Department of Computer Science
Connecticut College
New London, CT, USA
jnash1@conncoll.edu, parker@conncoll.edu, joconno2@conncoll.edu

***Abstract*— In order for evolution to populate the planet with multiple species, two processes need to be at work. One is speciation, which involves the development of a new reproductively isolated species from an ancestral one. The other is that a reproductively isolated species can evolve to be more complex and potentially more capable over time. This second process is what we address in the research reported in this paper. One of the issues with developing individuals with more complex attributes is that those attributes typically take multiple mutations to be viable and before that point, the changes made by mutations are often detrimental. In this research, we use a simulated environment filled with simple agents and a genetic algorithm operating on these agents, each of which has its own set of chromosomes. We use this to test the plausibility of the species developing more complex structures (in this case sight) as the genetic algorithm population survives and evolves over 1000s of generations. In order to better simulate intermediate detrimental mutations, each mutation results in an increased metabolic cost for the agent, leading to higher energy expenditure with every movement or action taken. Tests were done with differing number of mutations required before sight was developed. The results showed that evolution of complex attributes are possible depending on the number of detrimental mutations required to make the attribute an advantage to the agent.**

***Keywords*— Speciation, Evolution, Macroevolution, Genetic Algorithm, Fitness Valley**

I. Introduction

The theory of evolution requires that new species can develop from previous species and that species can evolve to have increasingly complex attributes. The goal of our research is to use Genetic Algorithms (GAs) to model evolution, so both of these requirements have to be addressed. In previous research, it was shown that allopatric speciation was possible in an environment of simple agents interacting as they moved, ate food, and mated with each other [1, 2]. We are using the latest of the models used in this research to experiment with the evolution of more complex attributes in a population of a single species. One of the issues with biological evolution is that complex structures such as wings and eyes cannot form through a single mutation. Several mutations are required, and in many cases, these mutations result in the agent being less fit than agents without one or more of these incremental mutations. We refer to this temporary lowering of an organism's fitness while evolving a complex organ as a fitness valley [3]. For example, the first step in evolving an eye would be to have an area of skin sensitive to light. Although the evolution of an eye may be advantageous and increase the fitness of the organism in its environment, traversing through the fitness valley in this case is probably also a liability since this area of skin is more vulnerable to injury. In the case of a wing, the first notable mutation would probably be a bump on the thorax, which could eventually develop through mutations in further generations to become a wing. However, this proto-wing would most likely hinder the fitness of the organism until further mutations were able to improve the use or efficacy of the wing.

Over time, several models have been developed to address the mechanisms of evolution previously discussed. One of the earliest approaches, Tierra, introduced a system of self-replicating digital organisms that compete for computational resources [4]. This system notably involved no exogenous fitness function, allowing the organisms in the environment to demonstrate rudimentary concepts of artificial evolution *in silico*. Shortly after, Avida expanded on this concept, providing a more structured framework for studying adaptive evolution through controlled experiments [5]. These models primarily focused on microevolutionary processes, emphasizing mutation and selection within relatively simple environments, and were highly effective at demonstrating adaptation but limited in their ability to explore complex traits that require overcoming fitness valleys. Following these foundational models, the Tangled Nature Model introduced a more ecological perspective by simulating evolutionary dynamics within complex adaptive systems, allowing for the study of species interactions and the role of macroevolutionary processes in shaping biodiversity [6]. This model laid the groundwork for understanding how populations traverse adaptive landscapes with multiple peaks and valleys, but like its predecessors, it did not focus on the evolution of complex, multi-step traits within individual organisms. More recent advancements, such as Aevol, extended the capabilities of these early systems by incorporating modular genomes, enabling the evolution of more sophisticated phenotypes [7]. This model was one of the first to investigate how genetic regulatory networks could facilitate the traversal of fitness valleys, making it possible for digital organisms to evolve traits with multiple interacting components. Building on these efforts, LaBar & Adami examined how complex traits could evolve in rugged fitness landscapes, focusing on genetic interactions and environmental variability [8]. Similarly, Arthur et al. applied digital evolution to investigate how functional capabilities arise and contribute to speciation under various selection pressures, demonstrating how small genetic changes can accumulate to form complex traits [9]. Despite these advancements, most of these models do not specifically address the evolution of complex organs that require overcoming fitness gaps through incremental mutations, as the algorithms were often geared towards optimizing simpler adaptive traits. Our research aims to address this gap by using GAs to simulate the effects of the step-by-step development of complex attributes, such as sight, within a controlled environment.

In our research, we use the environment previously developed to investigate the evolution of “sight” in digital organisms. While earlier versions of this environment only allowed agents to detect adjacent spaces, our experiments focus on agents evolving to sense two squares away from their position. This progression towards increased sensory capacity requires multiple mutations, each of which incurs an energy cost, simulating the incremental and often detrimental nature of intermediate evolutionary steps. This is not to say that we develop the mechanism for increased complexity/information to enter the chromosome. How that happens is still an open question. What we do is create a mechanism within the existing chromosome for mutations to create the stepping-stones for advanced features. We conducted a series of tests to determine how varying the number of required mutations and their associated energy costs impacted the likelihood of evolving sight. The results demonstrate that under the right conditions, even multi-step adaptations with an associated cost can evolve and become advantageous, confirming that complex traits can emerge depending on the evolutionary pressures and the structure of the fitness landscape.

As the learning system used in our environment, GAs are designed to emulate natural selection and have traditionally been used to find optimal arrangements of gene sequences to solve specific computational problems [10]. However, most standard GAs are limited in their ability to simulate the gradual emergence of new complexities or morphological traits in evolving agents. Some approaches, like NeuroEvolution of Augmenting Topologies (NEAT), introduced by Stanley and Miikkulainen, address the evolution of complex neural architectures that can enhance an agent’s control systems [11]. NEAT allows for the incremental evolution of increasingly sophisticated neural networks, utilizing genetic crossover and mutation to add new nodes and connections. This work has made significant strides in creating advanced control systems, but its primary focus is on neural evolution rather than the development of new physical attributes or morphological complexity.

Similarly, Genetic Programming (GP), introduced by John Koza, evolves tree-like structures to solve problems by modifying the programmatic logic of agents [12]. While GP has been successful in evolving complex, hierarchical solutions, it does not inherently support the stepwise formation of physical traits that require overcoming fitness valleys. Each of these systems introduces mechanisms for increasing control complexity but lacks a structured way to evolve complex traits that have an initially detrimental effect on an agent’s fitness. In our work, we address this limitation by focusing on the physical evolution of a sensory organ, demonstrating how complex attributes can emerge through incremental mutations, despite the associated temporary reduction in fitness.

This work is based on a model previously created by Parker and Nash [2], which is being rewritten for release as a standalone research environment. However, the Java source code for the model is currently available upon request.

II. ENVIRONMENT

The model uses a grid based representation of the environment, over which agents can move, eat, and interact with each other. This representation was chosen as it is simple enough to minimize confounding factors outside agent’s direct control, yet complex enough such that evolution could occur in a natural way. Additionally, using a grid allows us to easily modify the size of the environment and observe if results are consistent over different environment sizes.

The grid consists of discrete blocks (spaces) which can either be empty or occupied by some entity. In the research reported in this paper, the spaces can be occupied by either food (seeds) or the agents themselves. Seeds are randomly generated in the grid as time progresses, and act as a food source for the agents. These seeds can be of various sizes (small, medium, large) and are represented visually as squares of varying sizes. Agents are initialized randomly and are represented by circles of varying size and colors. The color is defined by the agent's genome and the size is defined by the size of the agent, which is inherited separately from the genome.

The environment progresses based on time cycles (turns). During each turn a single action is performed by each living agent. Additionally, food is generated randomly on the grid each turn. The number of turns passed from the initial state of the grid is said to be the age of the environment, similarly the number of turns an agent has survived is said to be the age of that agent. Visual examples of this environment can be found in previous work [2].

III. AGENTS

Agents are the evolving entities in our environment and the only entity on the grid that is able to move between spaces. The only restriction placed on their movement is that they cannot move to a space already occupied by another agent. If two agents are adjacent to each other, they may interact. In the current model, the only interaction available between agents is reproduction. As the agents may reproduce freely, the total number of agents present on the grid is variable as it may increase or decrease depending on the actions of the agents, the amount of seeds available, and the current population size.

The most impactful trait of an agent is its size, as this dictates the type of seed the agent is best suited to consume. Large agents prefer larger seeds, whereas smaller agents prefer smaller seeds. Thus, the size of an agent directly affects how well it will perform in the environment.

Seeds provide the agents with energy when eaten, the amount of energy depending on both the size of the agent and the size of the seed [2]. Agents also lose a certain amount of energy each turn, the amount lost being equal to the square root of their age. The amount of energy an agent loses each turn can be modified and such modifications to the energy burn rate are discussed later in this paper.

An agent starts with 100 initial energy. As the agent performs actions, it burns energy passively. The only other source of energy consumption is reproduction, which carries an additional cost of 80 energy. This cost is only levied when reproduction results in the production of a child agent, since if both parent agents agree to reproduce there is no chance of failure. Conversely, if either parent decides not to reproduce or does not have enough energy to reproduce, then there is no chance of reproduction. As the agents age, energy consumption increases, meaning that agents must reproduce in order to maintain a stable population. Older agents are unable to consume enough food to cover their energy burn, even if they eat a seed on each turn. Therefore, while agents have variable lifespans, there is an upper limit to how long an agent can survive.

Other agent traits, such as the agent controller, secondary traits, and reproductive preferences, are stored in chromosomes. An action chromosome informs the agent controller, discussed later, of the priorities for that agent. A reproductive choice chromosome determines possible mates for that agent, such as preferred color, age, and phenotype.

Notably, the reproductive choice chromosome does not consider the size of a possible mate. This is because size directly affects fitness, and the agents should choose mates based on their own preferences rather than a pre-defined fitness. The phenotype chromosome contains a series of traits that are solely used for determining if an agent is an acceptable mate by the reproductive chromosome, as well as the sight gene. The specifics of the sight gene are discussed in that respective section.

IV. Agent Sight

During the initialization of the environment, all starting agents are created with a sight range of 1. Therefore, these agents are only able to see spaces directly adjacent to them. These initial agents have their sight gene set to contain only 0 bits, representing a base state before any evolution has occurred. Over time, the agents may mutate during reproduction and gradually change these 0 bits to 1. As more bits in an agent's sight gene become 1, indicating an intermediate internal body change, more energy is required for the agent to move, so it loses more energy each turn. The functional change in an agent's sight is shown in Fig 1.

The increase in energy paid per bit, set as a hyperparameter before each test, scales linearly with the number of bits that become 1. This energy cost represents the cost of intermediate detrimental mutations that are often required in nature as more powerful senses are developed. In addition, more powerful eyes consume more calories and thus consume more of an organism's energy.

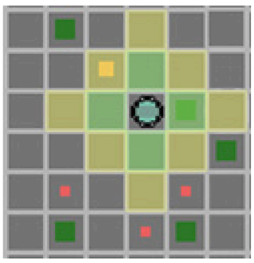

Fig. 1. Agents initially have a sight of 1, denoted by the green tinted squares. Once the agent evolves to attain a sight of 2, it is able to see the yellow tinted squares. If this medium agent had evolved sight, it's best move would be to move left or up, in order to get closer to the medium seed. If it had not evolved sight, it would likely move right, in order to consume the large seed, regardless of the lesser energy payoff.

If an agent has a sight gene containing all 1s, then that agent has evolved to have a sight range of 2. These agents are able to see spaces up to 2 blocks away from them, using blockwise distance. An agent does not benefit from having a sight gene containing a mix of 0s and 1s, all of the sight bits must be activated for the agent's sight to increase. Agents must accept some cost as they evolve sight, as they do not develop sight or receive any advantage after one random mutation. Instead, sight requires multiple mutations, each with a cost, evolved over many generations until full sight is achieved.

V. Agent Controller

Each turn, the agent controller will use the values in an agent's action chromosome to determine the actions of that agent during that turn. Agents have access to a variety of actions although some actions may be unavailable. During each turn, the immediate surroundings of an agent are analyzed to determine the action the agent should take.

The method for determining an agent's action differs from previous work using a similar model, as with the addition of sight the previous controller is no longer sufficient. Previously, each agent would try every action starting from the action with the highest priority. Once an action succeeded, the agent had finished its turn. However, with the addition of sight the agents must also consider direction, as if a high priority action could be taken after a turn of movement, it should begin moving in that direction. An example of this is shown in Fig. 2.

Instead of trying actions from highest to lowest priority, the agent observes the spaces around it (to a distance equal to the sight length of that agent) and determines a direction with the highest priority. This direction does not mean that a high priority action is immediately available in that direction, only that choosing that direction would either result in a high priority action or moving closer to a high priority action.

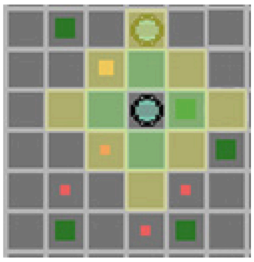

Fig. 2. This agent would consider the upwards direction to contain a medium seed and a medium agent, the leftwards direction to contain a medium seed and a small seed, with the other directions not containing items of as high interest to a medium agent. If the agent highly prioritized reproduction, it would assign a higher priority to the upwards direction than the leftwards direction, as there is another medium agent in this direction, and use it's turn to step upwards.

Once the agent chooses a direction, it then attempts to move in that direction. If there is another agent blocking that move, the agent then determines if it is willing to reproduce with that agent, attempting reproduction if so. If the move is successful, the agent then determines if there is a seed present on its new space. If so, the agent determines if it should eat that seed or not before its turn ends, if no seed is present the agent's turn ends immediately.

VI. GENETIC ALGORITHM

The priorities assigned by the agents to each action, the reproductive preferences, the agent's phenotypes, and the sight gene for each agent are all specified and evolved by a GA. The agents pass down their characteristics to their children via genetic operations on the parents' chromosomes.

The selection process is entirely determined by the agents instead of a fitness function; if two agents are adjacent to each other and their reproductive preferences are met by the other agent, then they are able to reproduce. The idea of selection based on agent fitness does not fully apply here, as any agent can reproduce with any other agent. A more fit agent that consumes more seeds is more likely to live longer and thus reproduce more, but this is not a guarantee.

The GA used in our model is most similar to a steady state GA, except that we do not use standard chromosome replacement, nor do we have a fixed population. Instead, old individuals die when out of energy and new individuals are added to the population via agent reproduction without regard to the overall population size. The population does have a soft maximum, as only so much food is generated each turn, but that is the only major limiting factor on the size of the population.

VII. RESULTS

Following the testing methodology of previous work, initial tests were done to check the model [1]. These initial tests showed that the agent populations could evolve, optimize for the available food sources, and sustain a population only limited by the available number of seeds. During these initial tests the evolution of sight was disabled, and the agents were limited to a sensor distance of 1 at all times.

The main tests were designed to test the ability of the population to evolve sight and observe the effects of this evolution on the overall population. During these tests, seeds of random sizes were created throughout the grid and the population was allowed to inhabit the grid without obstacles or other considerations. The population was created by generating a random number of agents with random sizes and chromosomes. Initially, all agents started with their sight gene containing all 0s, resulting in a sensor distance of 1. During each test the model was run for 500,000 turns.

During the first test, the sight gene was set to contain 2 bits such that both bits would have to evolve to be 1 before the agent obtained sight. In the first test, the cost for each sight bit of 1 was set at a 10% increase in energy consumption. These parameters were tested in 10 different trials and the agent population developed sight during every trial. The percentage of the population possessing increased sight rapidly climbed to 100% after sight initially evolved, however due to the increased energy cost this also resulted in an increased the average number of seeds consumed per agent and a drop in the total population and average number of children per agent. Fig. 3 shows the results of one of these trials.

In the second test, the number of sight bits was increased to 5, with the penalty per active sight bit being reduced to 2.5%. With these parameters, the agent population developed sight once over the course of 20 trials. The effects of this evolution were largely similar to the effects during the first test, with decreased overall population and increased individual seed consumption. Fig. 4 shows the results of the successful trial with these parameters.

The third test increased the number of sight bits to 10 and the penalty reduced further to 1%. The agent population was unable to successfully evolve sight in any of the 40 trials performed. As an extension to this testing, these parameters were tested with a total model runtime of 2,000,000 turns for 5 trials. Even with this increased runtime, the agents did not evolve sight in any of our trials. We believe that given enough time or enough trials, the agents may still evolve sight under these conditions, even with the complexity of this evolution being much higher.

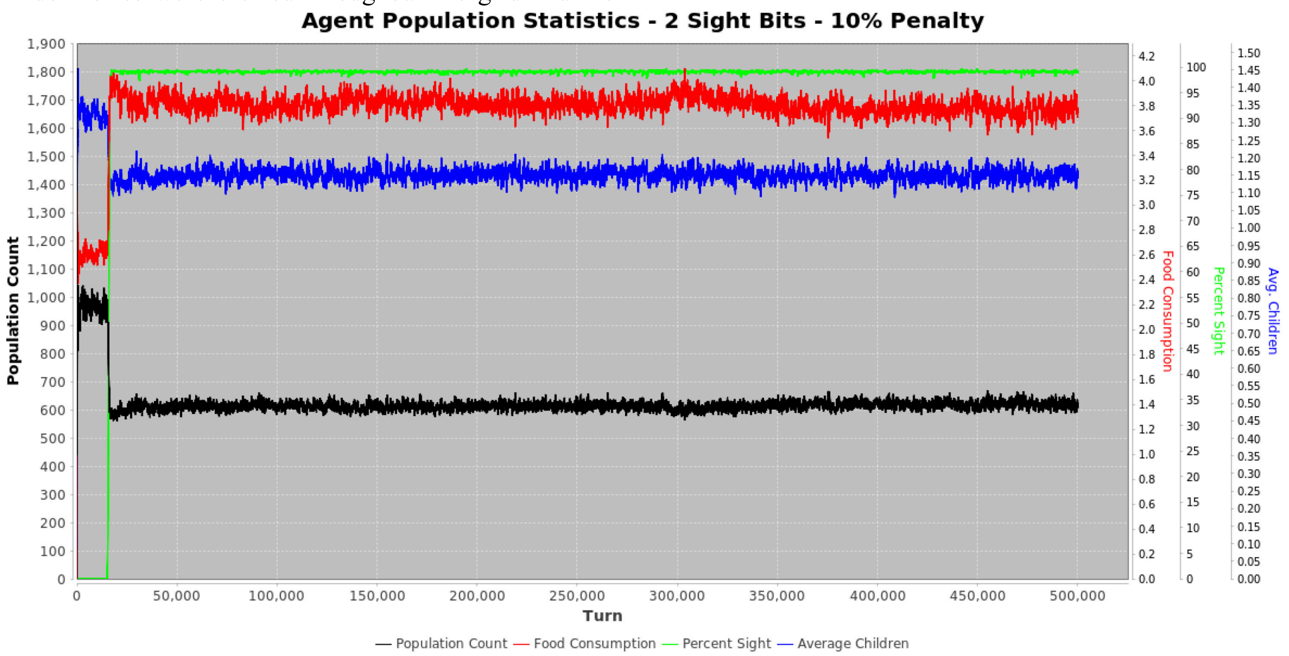

Fig 3. Graph of agent population over time during the first test. During this test the agents had a total of 2 sight bits and received a 10% penalty for each active sight bit.

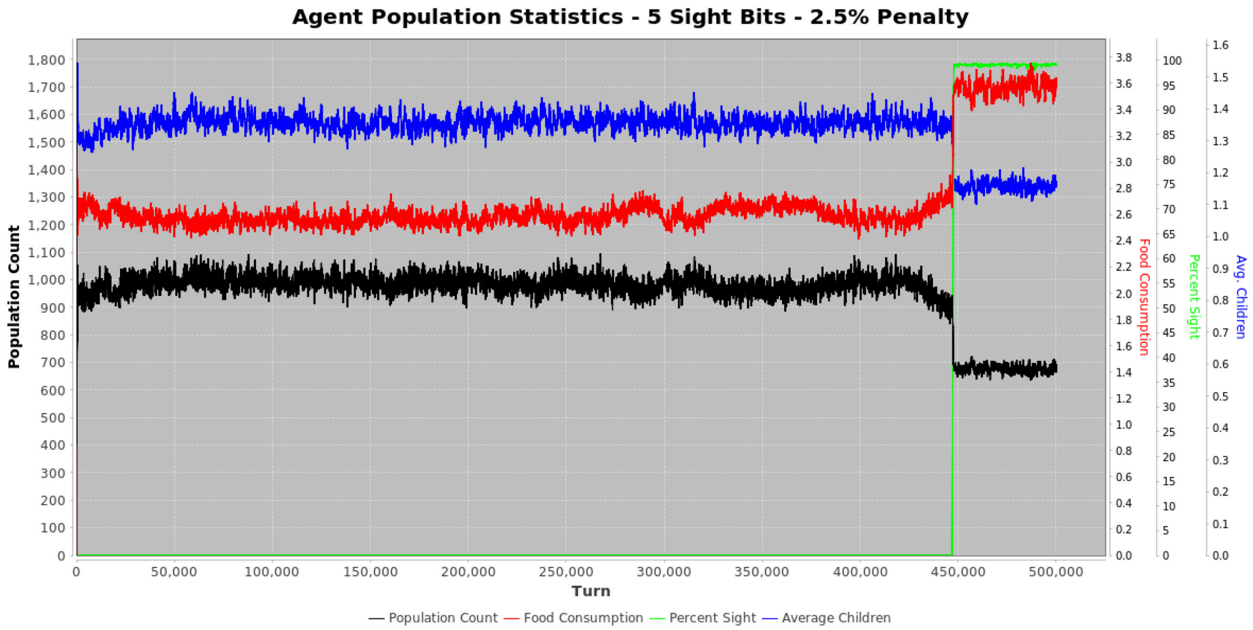

Fig 4. Graph of the agent population that successfully developed sight over time during the 16th trial of the second test. During this test the agents had a total of 5 sight bits and received a 2.5% penalty for each active sight bit.

It is interesting to consider the results of these three tests. In the first test, attaining sight required the mutation of 2 bits and a total detrimental effect of 20% in increased energy used per move, and sight evolved quickly. In the second test, 5 mutations and a total detrimental effect of 12.5% was required, and sight evolved much slower. In the third test, 10 mutations and a total detrimental effect of 10% was required, and sight did not evolve even with additional turns. This implies that the number of mutations required is more of a factor than the total detrimental effect.

VIII. CONCLUSION

A key part of evolution is the ability to evolve new capabilities and traits, even if the intermediate steps to a positive result are detrimental. Our model has shown that the simulation of this aspect of evolution is possible given enough time to overcome the negative evolutionary pressure caused by the detrimental intermediate stages.

Additionally, our model does not rely on fitness to determine agent reproduction. The agents themselves determine their reproductive partners, without external pressure applied to encourage any specific choice or direction of evolution. Standard GAs display this behavior when overcoming local minima, but require some type of external fitness function to correctly optimize. In contrast, our model overcame the local minima of reduced sight and pushed through negative feedback to evolve increased sight without information regarding the fitness of any individual agent.

In terms of the overall population, the evolution of sight is clearly detrimental as it causes a higher consumption of a limited common resource. However, increased sight is highly beneficial for an individual, as it provides a competitive advantage over agents without increased sight and is necessary to compete with other agents that possess increased sight.

In future work, we would like to explore the idea of certain evolutionary choices having positive effects on individuals yet negative effects on the overall population. The addition of the possibility of teamwork, such as allowing agents to decide to share some energy with other agents, could align the evolutionary process more closely with the interests of the overall population. We also plan to use these results to improve our work in simulating speciation, the divergence of a parent species into multiple distinct new species, in order to create a more complete simulation of evolutionary processes.

REFERENCES

[1] G. B. Parker and T. B. Edwards, "Using a Genetic Algorithm to Replicate Allopatric Speciation," 2019 IEEE Congress on Evolutionary Computation (CEC), Wellington, New Zealand, 2019

[2] Parker, G. and Nash, J. (2024). Using Secondary Inherited Characteristics During Reproductive Choice to Replicate Allopatric Speciation. In *Proceedings of the 16th International Joint Conference on Computational Intelligence – ECTA*.

[3] C. S. Gokhale, Y. Iwasa, M. A. Nowak, and A. Traulsen, "The pace of evolution across fitness valleys," Journal of Theoretical Biology, vol. 259, no. 3, pp. 613–620, Aug. 2009

[4] Ray, T. S. 1991. An approach to the synthesis of life. In : Langton, C., C. Taylor, J. D. Farmer, & S. Rasmussen [eds], Artificial Life II, Santa Fe Institute Studies in the Sciences of Complexity, vol. XI, 371–408. Redwood City, CA: Addison-Wesley.

[5] C. Adami and C. Titus Brown, "Evolutionary Learning in the 2D Artificial Life System 'Avida,'" Artificial Life IV. The MIT Press, pp. 373–377, Jan. 1994

[6] K. Christensen, S. A. Di collobiano, M. Hall, and H. J. Jensen, "Tangled Nature: A Model of Evolutionary Ecology," Journal of Theoretical Biology, vol. 216, no. 1, pp. 73–84, May 2002

[7] B. Batut, D. P. Parsons, S. Fischer, G. Beslon, and C. Knibbe, "In silico experimental evolution: a tool to test evolutionary scenarios," BMC Bioinformatics, vol. 14, no. Suppl 15, p. S11, 2013

[8] T. LaBar and C. Adami, "Different Evolutionary Paths to Complexity for Small and Large Populations of Digital Organisms," PLOS Computational Biology, vol. 12, no. 12, p. e1005066, Dec. 2016

[9] R. Arthur, A. Nicholson, P. Sibani, and M. Christensen, "The Tangled Nature Model for organizational ecology," Computational and Mathematical Organization Theory, vol. 23, no. 1, pp. 1–31, Feb. 2016

[10] J. H. Holland, Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. UMichigan Press, 1975.

[11] K. O. Stanley and R. Miikkulainen, "Evolving Neural Networks through Augmenting Topologies," Evolutionary Computation, vol. 10, no. 2, pp. 99–127, Jun. 2002.

[12] J. R. Koza, Genetic programming. Cambridge, Mass: Mit Press, 1998.