

Harnessing Evolution: The Interaction Between Sexual Recombination and Coevolution in Digital Organisms

by Mairin Chesney

“So far, we have been able to study only one evolving system, and we cannot wait for interstellar flight to provide us with a second. If we want to discover generalizations about evolving systems, we will have to look at artificial ones.”

J. M. Smith [14]

Problem and Motivation

As the complexity of software systems increase, they become harder and harder to design. One of the most complex man-made systems is IBM’s Watson, a prime example of research to engineer true intelligence. The only notably more successful instances of intelligence were produced by evolution, with the human brain being at the very top of that scale, both for robust intelligence and sheer complexity. Evolution is clearly a powerful constructive force, but fundamentally it is just a natural algorithmic process. Unfortunately, evolution is hard to study, with scientists typically gleaning information from past events such as observing the fossil record, which is both low resolution and full of gaps. In order to study evolution as it happens, simple systems such as *Escherichia coli* and *Drosophila melanogaster* are often used. However, even in these carefully controlled laboratory settings, observing each individual is impossible, the timescales required are still on the order of years for individual experiments, and the applicability for understanding the evolution of intelligence is distant.

Digital evolution is an alternative method for understanding evolutionary dynamics that bridges the gap between pure biological systems and applied evolutionary computation. Whereas in traditional evolutionary computation, the user specifies an explicit fitness function based on a fixed goal, digital evolution is more open-ended. Fitnesses of digital organisms are based on efficient resource utilization leading to rapid reproduction. A “fit” digital organism propagates more quickly than a less fit organism and is able to gain a greater foothold in the environment.

An evolutionary study in a purely biological system can consume decades [6]. Digital organisms offer numerous advantages over traditional methods of studying evolution including speed and efficiency [7]. The gaps in the fossil record that haunt scientists and anthropologists are nonexistent when using digital organisms. An organism’s entire ancestral line can be saved, and each mutation can be isolated and analyzed in the environment in which it occurred, as well as in different environments. Digital evolution is not a replacement for biological systems, rather, it is an extraordinary compliment. If a biologist has a hypothesis and wishes to test its validity, he or she can do so in a digital system first, potentially saving years of work. Then, if something is learned about the evolution of a digital system, these results can also be applied towards developing new computational methods.

Using digital evolution, we are exploring the interaction between parasite-host coevolution and the evolution of sexual recombination. The origin and maintenance of sexual recombination is one of the least understood mysteries related to evolution [10], but also one of the most important for understanding how evolution produces complex solutions to problems. Sex comes with many challenges: finding a suitable mate, avoiding sexually-transmitted diseases, maintaining the physiological mechanisms to recombine genomes, and hoping the remixed genomes produce a sensible result. The greatest challenge may be the two-fold cost of sex, which John Maynard Smith described from the standpoint of a female individual. As a sexual organism, she puts only half of her DNA into an offspring, but if she had switched to be asexual she could have put all of her DNA in her offspring, doubling the amount of her genetic material that makes it into the next generation [15]. Expanding this view to the population level, since males do not directly birth offspring, the population as a whole could produce twice as many offspring if everyone were asexual.

The origins of host-parasite coevolution are better understood, but many questions still exist about the role of these dynamics in adaptive evolution, particularly in regards to the origins and evolution of sex.

In *Through the Looking Glass* by Lewis Carroll, the Red Queen says to Alice, “It takes all the running you can do, to keep in the same place” [1]. From this remark comes the term “Red Queen dynamics”, which are said to occur when competing species pressure each other to continually evolve in order to maintain the same relative fitness. In the context of host-parasite coevolution, hosts must continually evolve away from parasites or else they will be taken over by them. Likewise, parasites must continually evolve to be able to better target hosts or else the hosts will escape from them. As hosts and parasites evolve, what is fit for either hosts or parasites in one generation may quickly become unfit, thus favoring mechanisms to accelerate evolution, such as genetic recombination [13]. Indeed, the pressure for heightened evolvability due to Red Queen dynamics is one of the most widely accepted theories regarding the origin and maintenance of sex.

Our goal here is to perform experiments with digital organisms to understand how Red Queen dynamics may have produced sexual recombination and how sex continues to be maintained. Further, our longer-term goal is to apply these principles to achieve the levels of complexity necessary for artificial intelligence. Nearly all complex life is sexual (or at least went through a sexual phase in its history) implying that sex is very important for the evolution of complex traits. If we can understand how and why sexual recombination exists, we will simultaneously solve a huge mystery in biology and provide computer science with a more refined tool for developing novel and complex solutions to computational problems.

Background and Related Work

An area of computer science with a demonstrated record of producing human-competitive results (including patents and real-world successes) is genetic programming [5]. Genetic programming draws its inspiration from biological evolution. However, it fails to meet the qualifications of true Darwinian evolution. Although it is an evolutionary system, reproductive success is decided by how well a user-defined problem is solved. Digital evolution, on the other hand, allows us to explore natural evolutionary processes since it displays the defining characteristic of Darwinian evolution: heritable variation of reproductive success.

The earliest ancestors of digital organisms were the rampant computer viruses of the 1980s [3]. Although these viruses self-replicated, they did not mutate and thus could not improve. Mutation of computer programs was first introduced by Steen Rasmussen [11] in a simulated world built upon the Core War game [2], in which programs competed in a virtual battleground for computer resources. Early digital organisms were too unstable for much scientific study. Soon after, Thomas S. Ray created the system Tierra and performed the first successful experiments with evolving populations of computer programs [12]. In 1993, Christoph Adami, Charles Ofria and C. Titus Brown created the artificial life platform Avida, extending these techniques into a proper scientific research platform suitable for replicated and well-controlled experiments [9].

Avida is an example of the type of artificial system John Maynard Smith was referring to in his quote at the beginning of this paper. Avida is an artificial life software platform that consists of populations of self-replicating digital organisms. The populations occupy a virtual environment with resources in the form of virtual CPU cycles (their form of energy). Each organism attempts to gain use of these resources by performing a simple computational task, such as adding two numbers together. Instead of consisting of strands of DNA, digital organisms in Avida (Avidians) consist of lines of computer instructions. Each organism makes up a short, syntactically correct program. Avidians have a virtual CPU used to execute their genome, which requires CPU cycles. Variability in organisms is introduced as they replicate; mutations can occur on any instruction when it is copied or during the birth of an offspring. An instruction might be added, deleted, or substituted for another instruction. As organisms mutate, some may break or reproduce more slowly, some occasionally reproduce more quickly. This variation allows natural selection to drive true Darwinian evolution.

Unlike typical digital organisms, digital parasites do not receive CPU cycles directly and must therefore infect a host. Instead of *dividing*, parasites *inject* their offspring into a host. In our setup, parasites have nine vectors by which they can infect hosts. Each vector is associated with a computational task (in this case, one- or two-input bitwise Boolean logic operations). A parasite must perform one of these tasks to infect a host, but the host can block the parasite by performing the same task. If a parasite is successful, it is

treated like a high-priority thread in the host organism, stealing CPU cycles. A parasite may steal only a few CPU cycles and sicken the host slightly, or it may steal all of the host's CPU cycles, effectively killing the host. Thus, in order to fully avoid infection, a host must perform all nine tasks that potential parasites target. However, organisms must evolve these tasks, and once they do, they must spend CPU cycles on performing the tasks. Therefore, increased resistance comes at a cost. Organisms should perform tasks to avoid common parasites, but must avoid performing such a large number that it becomes more computationally expensive than infection.

The only variation that occurs in asexual organisms is a result of mutations. Each resistance trait must be individually evolved, and varying skills acquired by different organisms cannot be shared. Sex allows rapid mixing and matching of genetic code potentially performing different tasks, which maintains diversity and may facilitate the evolution of host resistance. Sexual organisms can evolve resistance, recombine, and share their abilities. In Avida, two sexual organisms recombine by swapping a continuous region of code and the two resulting offspring are placed back in the population [4]. When the organisms swap chunks of code, they are able to share their abilities. We hypothesize that this sharing of abilities gives sexual organisms an edge over their asexual brethren.

Approach and Uniqueness

The interaction between host-parasite coevolution and the evolution of sexual recombination has been explored from a biological standpoint [16]. Our approach harnesses the power of digital evolution to explore if parasites may have driven organisms to evolve sexual recombination over asexual recombination. By applying the benefits of both evolutionary biology and computer science, the understanding we gain about the interaction between parasites and recombination can be applied towards both fields. To fully explore the interactions between host-parasite coevolution and the evolution of sexual recombination in digital evolution, we decided to explore both directions of interaction - how parasites affect sex and how sex affects parasites.

We first looked at the effect of sexual recombination on parasite persistence. We completed one treatment of runs with obligately sexual organisms and one treatment of runs with obligately asexual organisms. Both treatments had parasites and consisted of 30 runs at 200,000 updates. Updates are the standard unit of time in Avida, and these runs corresponded to approximately 10,000 generations, but this measure varies substantially depending on the specific course evolution follows in a particular run. This set of runs allowed us to observe how the method of reproduction affected the parasites survival in the environment.

We then approached the question from a completely different angle. We again had two treatments of 30 runs at 200,000 updates, but this time one treatment had parasites and the other had no parasites. The organisms all started sexual, but their method of reproduction could evolve. By comparing organisms coevolved with parasites to organisms evolving alone, we were able to see how pressure from parasites affected the maintenance of sexual recombination.

For the first set of runs, we expected the sexual organisms to do a much better job at avoiding parasites than the asexual organisms. Asexual organisms must sit around and wait for mutations in order to develop increasing levels of parasite resistance. On the other hand, sexual organisms can rapidly mix and match their genome and share resistance abilities. Since sexual recombination likely drives higher levels of parasite resistance, we expected the populations in an environment with parasites to have much higher proportions of sexual organisms than the populations without parasites.

Results and Contributions

Our study explores both directions of host-parasite coevolution and the evolution of sexual recombination. We found from our first set of runs that obligately sexual organisms are much more successful at fending off parasite infection compared to their asexual counterparts (Figure 1). In sexual populations, the number of parasites grew rapidly while the organisms were initially unable to avoid them, but the hosts evolved high levels of resistance, and the parasite count dropped quickly. In asexual

populations, the number of parasites grew rapidly and stayed high before slowly beginning to drop near the end of the experiment. These results are consistent with our expectations. When dealing with a new parasite, both sexual and asexual organisms eventually develop resistance. However, sexual organisms are much better equipped to rapidly fend off parasites than asexual organisms.

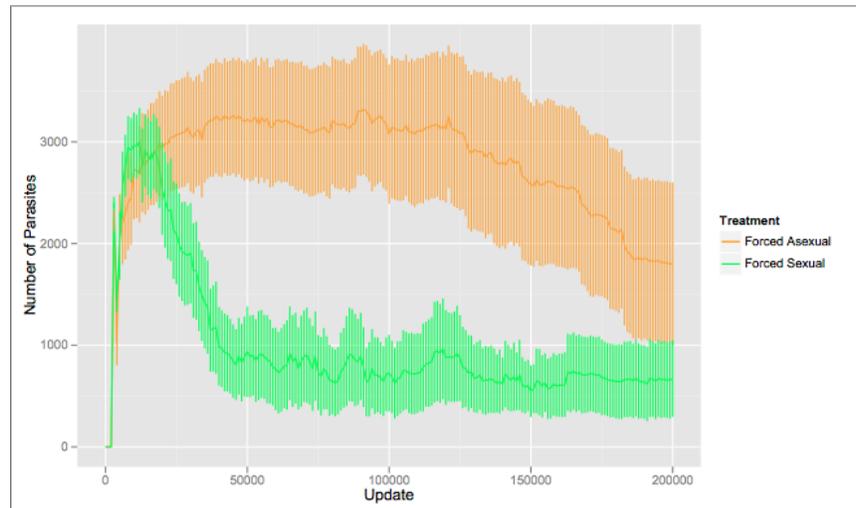


Figure 1: Average number of parasites in populations of obligately asexual organisms (orange) compared to populations of obligately sexual organisms (green) over time.

Although the results of the first set of runs were nearly as expected, the second set of runs confounded our expectations. When given the choice between asexual and sexual reproduction, hosts in populations with parasites seemed to switch from sexual recombination to asexual recombination slightly more quickly than hosts in populations with no parasites (Figure 2). We do not yet see that sexual organisms evolve preferentially over asexual organisms in the presence of parasites. Since sexual recombination gives hosts such a distinct advantage over parasites, why would they fail to fix it in a population? A number of possibilities contribute to these results.

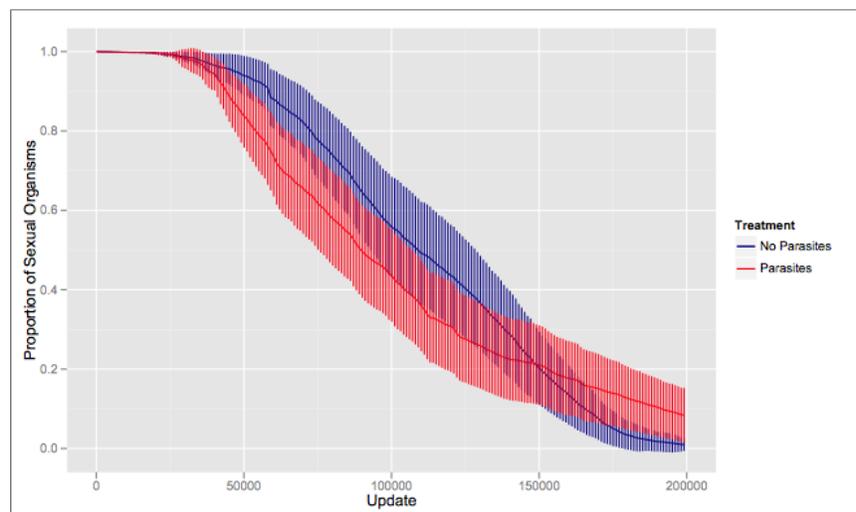


Figure 2: Average proportion of sexual organisms in an environment with parasites (red) compared to no parasites (blue) over time.

The first possibility is that sexual digital organisms have an implicit waiting time that allows asexual organisms to spread throughout the environment more rapidly. Sexual organisms must wait for a partner in

order for them to recombine. We tested this idea by forcing asexual organisms to wait an equivalent amount of time, but we found no change in our results. Another possibility is that Avida's version of sexual recombination is extraordinarily harsh on organisms. Because insertion and deletion mutations can change an organism's length, swapping chunks of code from two separate programs can cause a lot of damage. One idea to counteract some of the harm done is to restrict organism length. If two sexual organisms are the same length, they are theoretically less likely to break upon recombination. We are currently exploring this hypothesis and brainstorming other ideas.

Digital evolution presents a number of interesting computational challenges. Life in the natural world is inherently complicated. The number of interactions between genes, organisms, and populations is extraordinarily large and rarely fully understood. Translating even a small piece of these interactions into a computational environment requires careful consideration. Even within our "simple" Avida model, the modifications are still endless. But when studying a biological theory, the temptation to modify without any biological backing must be resisted. Finally, the required computer memory is enormous. I was put on our lab's "list of shame" for generating over 60 gigabytes of data in a single set of runs.

The study of evolutionary theories using digital evolution has important fundamental intellectual merit. We are born with the desire to answer the question, "Why?" Digital evolution provides insight into the evolution of complex traits, and can be easily manipulated to study hugely varying characteristics of biology such as cooperation, sexual reproduction, predator-prey interactions, and host-parasite relationships. As we gain understanding of how and why organisms evolve one way or another, we may start understanding the directions organisms will evolve in the future. We learn about making use of evolutionary processes as we understand them further. In hopes of achieving more and more complex solutions using evolutionary computation, we strive to understand how a trait as complex as sexual recombination evolves, and how its presence may have driven organisms toward greater complexity and intelligence. If parasites have the power to drive organisms to a novel form of reproduction, their abilities can be harnessed by evolutionary computation to solve real-life problems where traditional evolutionary computation fails [8].

By harnessing the power of both biological and computational systems, we are able to use an interdisciplinary approach to explore evolutionary theories. If a biological theory is supported in a computational system, the results benefit both fields. Biologists learn how organisms live and interact, and computer scientists learn what natural phenomena can help develop newer and better methods of solving problems. If we can figure out a method by which digital organisms achieve greater levels of complexity, these findings can be applied towards new approaches to genetic programming and artificial intelligence.

References

1. Carroll, L. (1871). Through the Looking-Glass and What Alice Found There.
2. Cohen, F. (1987). Computer viruses: theory and experiments. *Computers & security*, 6, 22-35.
3. Fortuna, M. A., Zaman, L., Wagner, A. P., & Ofria, C. (2013). Evolving Digital Ecological Networks. *PLOS Computational Biology*, 9, e1002928.
4. Frénoy, A., Misevic, D., Taddei, F. (2011). Digital Sex: Causes and Consequences. In *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*. (pp. 532-533). MIT Press.
5. Koza, J. R., Bennett III, F. H., & Stiffelman, O. (1999). *Genetic programming as a Darwinian invention machine* (pp. 93-108). Springer Berlin Heidelberg.
6. Lenski, R. E. (2013). Lessons Learned from 50,000 Generations of E. coli Evolution. *Annual Review of Microbiology*, 67.
7. Lenski, R. E., Ofria, C., Pennock, R. T., & Adami, C. (2003). The evolutionary origin of complex features. *Nature*, 423, 139-144.
8. McKinley, P., Cheng, B. H., Ofria, C., Knoester, D., Beckmann, B., & Goldsby, H. (2008). Harnessing digital evolution. *Computer*, 41, 54-63.
9. Ofria, C., & Wilke, C. O. (2004). Avida: A software platform for research in computational evolutionary biology. *Artificial life*, 10, 191-229.
10. Otto, S. P. (2009). The evolutionary enigma of sex. *The American Naturalist*, 174, S1-S14.
11. Rasmussen, S., Knudsen, C., Feldberg, R., & Hindsholm, M. (1990). The coreworld: Emergence and

- evolution of cooperative structures in a computational chemistry. *Physica D: Nonlinear Phenomena*, 42, 111-134.
12. Ray, T. S. (1991). An approach to the synthesis of life. *Artificial life II*, 10, 371-408.
 13. Salathé, M., Kouyos, R. D., & Bonhoeffer, S. (2008). The state of affairs in the kingdom of the Red Queen. *Trends in Ecology & Evolution*, 23, 439-445.
 14. Smith, J. M. (1992). Byte-sized evolution. *Nature*, 355, 772-773.
 15. Smith, J. M. (1993). *The theory of evolution*. Cambridge University Press.
 16. Zaman, L., Devangam, S., & Ofria, C. (2011). Rapid host-parasite coevolution drives the production and maintenance of diversity in digital organisms. In *Proceedings of the 13th annual conference on Genetic and evolutionary computation* (pp. 219-226). ACM.