

The Surprising Creativity of Digital Evolution

Joel Lehman¹, Jeff Clune^{1,2}, Dusan Misevic³

¹Uber AI Labs, San Francisco, CA, USA

²University of Wyoming, Laramie, WY, USA

³Center for Research and Interdisciplinarity, Paris, France

Corresponding author: Dusan Misevic, dule@alife.org

Abstract

Natural evolution and complex adaptations often surprise scientists. However, the creativity of evolution is not limited to the natural world, transcending any particular substrate. In the context of digital evolution, artificial organisms evolving in computational environments are also able to elicit surprise and wonder. Indeed, most digital evolution researchers can relate anecdotes highlighting how common it is for their algorithms to creatively subvert their expectations or intentions, expose unrecognized bugs in their code, produce unexpectedly potent adaptations, or engage in behaviors and outcomes uncannily convergent with ones found in nature. Such stories routinely reveal the surprise and creativity of evolution in these digital worlds, but they rarely fit into the standard scientific narrative and are treated as obstacles to be overcome rather than interesting results. Bugs are fixed, experiments are refocused, and one-off surprises become stories traded among researchers through lossy, inefficient and error-prone oral tradition. Moreover, to our knowledge, no collection of such anecdotes has been published before and many natural scientists do not recognize how lifelike digital organisms are and how natural their evolution can be. We have crowd-sourced the writing of a paper and collected first-hand reports from artificial life and evolutionary computation researchers, creating a written, fact-checked collection of entertaining and important stories. It serves to show that evolutionary surprise generalizes beyond the natural world, and may indeed be a universal property of all complex evolving systems.

Evolution provides countless examples of creative, surprising, and amazingly complex solutions to life's challenges, from flowers acting as acoustic beacons to parasites reprogramming host brains, constantly surprising even the seasoned field biologists. But the creativity of evolution, inventing something both original and effective (Runco, 2012), need not be constrained to the organic world. Independent of medium, evolution happens wherever replication, variation, and selection intersect. Thus, evolution can be instantiated digitally either to study evolutionary dynamics experimentally or to solve engineering challenges through directed digital breeding. Evolution as a whole can be considered creative, at least in the sense that it produces surprising and complex solutions that would be deemed as creative if produced by a human. We make a case that digital evolution experiments often produce strange, surprising, and creative results.

Most anecdotes about unexpected but singular results spread only through oral tradition, providing laughs and

discussion at professional conferences or comic relief during public talks. But they fail to inform the field as a whole in a lasting and stable way. To compile a comprehensive archive, we sent out a call for anecdotes to prominent digital evolution mailing lists and succeeded in reaching founders of the field as well as young, upcoming researchers. We then curated over 90 submissions to create a "greatest hits" collection, currently available as a pre-print (Lehman, 2018). In this extended abstract we summarize the full paper and include a small selection of its anecdotes. Besides an exercise in humility this provides insight and knowledge for practitioners, by showing the pervasiveness of obstacles and ways to overcome them.

We define surprise broadly as observing an outcome that significantly differs from expectations, whether those expectations arise from intuitions, predictions from past experiences, or from theoretical models. And at first, it may seem counter-intuitive that a class of algorithms can consistently surprise the researchers who designed them. From theoretical computer science we know that the outcome of a program often cannot be predicted without actually running it (Turing, 1936) and that simple programs can yield complex and surprising results (Runco, 2012). While we offer no formal proof of digital evolution's ability to generate surprise, the diversity of anecdotes highlights how common and widespread such surprising results are in practice.

The anecdotes collected roughly cluster into four representative categories: selection gone wild, in which digital evolution reveals the divergence between what an experimenter is asking of evolution and what they think they are asking; unintended debugging, in which digital evolution reveals and exploits previously unknown software or hardware bugs; exceeded expectations, in which digital evolution discovers solutions that exceed the expectations of the experimenter; and convergence with biology, in which digital evolution discovers solutions surprisingly convergent with those found in nature, despite vast divergence in medium and conditions. Here we briefly highlight representative anecdotes from each of these categories.

Why Walk When You Can Somersault? In a seminal work from 1994, Karl Sims evolved 3D virtual creatures that could discover walking, swimming, and jumping behaviors in simulated physical environments (Sims, 1994). However, when Sims initially attempted to evolve locomotion behaviors, things did not go smoothly. Instead of inventing clever limbs or snake-like motions that could push them along

(as was hoped for), the creatures evolved to become tall and rigid, so they would fall over, harnessing their initial potential energy and even performed somersaults to extend their horizontal velocity. Building on Sims' work, but using a different simulation platform, Krcah also discovered somersaulting without jumping at all (Krcah, 2008). Individuals with a few large blocks reminiscent of a "head" supported by a long thin vertical pole would "kick" the foot of the pole, falling head-first, somersaulting its foot, and scoring high on a jumping metric that measured a gap between the ground and the "lowest point" of the creature.

Floating Point Overflow Lands an Airplane. In 1997, Feldt applied digital evolution to mechanical simulations to evolve mechanisms that safely, but rapidly, decelerate aircraft as they land on an aircraft carrier (Feldt, 1997). Instead of slowly converging, evolution quickly produced nearly perfect solutions that were very efficiently braking the aircraft, even when simulating heavy bomber aircraft coming in to land. Indeed, the results were too good to be true: evolution discovered a loophole in the force calculation for when the aircraft's hook attaches, estimating force as zero when it is a number too large to store in memory and thus leading to a perfect score. Interestingly, insights from this experiment led to theories about using evolution in software testing (to find bugs and explore unusual behavior) and engineering (to help refine knowledge about requirements) (Feldt, 2002).

Parasites, Muscles and Bones. Maybe the most famous example of unintended biological convergence comes from the very first time Tierra system ran without crashing (Ray, 1992). It produced fascinating, complex ecologies, including parasitism, obligate sociality, and primitive forms of sexual recombination. More recently, in Cheney et al.'s virtual creatures (Cheney 2012), evolution generated locomotion strategies unexpectedly convergent with biological ones, by discovering scratch the benefit of complementary (opposing) muscle groups, similar to such muscle pairs in humans (biceps and triceps), combining them with stiff, bone-like material, resulting in a gait reminiscent of a horse's gallop.

Towards Complex Behavior via Temporary Setbacks. In a pioneering study Lenski et al. used the Avida platform to test Darwin's hypotheses about the evolution of complex features (Lenski, 2003). The researchers specifically focused on whether and how Avidians might evolve to perform the most complex logical function they were rewarded for. The most surprising outcome of all was that the pathway that evolution followed was not always an upward climb to greater fitness, but included some significantly deleterious mutations that reduced fitness up to two fold! These mutations set up a condition that allowed a subsequent beneficial mutation to complete the complex trait, just like traversing rugged fitness landscapes with deceptive fitness may happen during evolution in biological systems.

A persistent misunderstanding is that digital evolution cannot meaningfully inform biological knowledge because "it is only a simulation." By collecting and validating the true stories from the original scientists we directly show that these algorithms indeed unleash the creativity of the Darwinian process. Beyond biology, the ubiquity of surprising and creative outcomes connects to the nascent field of artificial intelligence safety: Many researchers therein are concerned with the potential for perverse outcomes from optimizing

reward functions that appear sensible on their surface. The list compiled here provides additional concrete examples of how difficult it is to anticipate the optimal behavior created and encouraged by a particular incentive scheme. Additionally, digital evolution may provide an interesting training ground for developing intuitions about incentives and optimization, to better ground theories about how to craft safer reward functions for AI agents.

Across a compendium of examples we have reviewed, there are many ways in which digital evolution produces surprising and creative solutions. The diversity and abundance of the collected examples suggest that surprise in digital evolution is a rule, rather than an exception. For every story we received or heard, there are many more that have been already forgotten as researchers retire. The ubiquity of anecdotes also suggests that creativity is not confined to evolution in nature, but may be a pervasive feature of evolutionary processes in general. While each case is surprising, these surprises are so frequent they can be considered routine. These anecdotes thus serve as evidence that evolution—whether biological or computational—is inherently creative, and should routinely be expected to surprise, delight, and even outwit us.

Note: Due to limited space here, we are unable to include over 50 co-authors and their affiliations and we refer the readers to the pre-print (Lehman, 2018) for this information.

References

- Cheney, N., MacCurdy, R., Clune, J., Lipson, H. (2013) Unshackling evolution: evolving soft robots with multiple materials and a powerful generative encoding. In GECCO Proceedings, 167–174.
- Feldt, R. (1998) Generating Diverse Software Versions with Genetic Programming: an Experimental Study. IEE Proceedings - Software Engineering, 145:228–236.
- Feldt R. Biomimetic software engineering techniques for dependability. Department of Computer Engineering, Chalmers University of Technology. Gothenburg, Sweden; 2002.
- Krcah, P. (2008) Towards Efficient Evolutionary Design of Autonomous Robots. In Proceeding Evolvable Systems: From Biology to Hardware ICES, 153–164.
- Lehman L, Clune, J., Misevic, D., Adami, C., Beaulieu, J., Bentley, P. J., Bernard, S., Beslon, G., Bryson, D. M., Chrabaszcz, P., Cheney, N., Cully, A., Doncieux, S., Dyer, F. C., Ellefsen, K. O., Feldt, R., Fischer, S., Forrest, S., Frénoy, A., Gagné, C., Le Goff, L., Grabowski, L. M., Hodjat, B., Hutter, F., Keller, L., Knibbe, C., Krcah, P., Lenski, R. E., Lipson, H., MacCurdy, R., Maestre, C., Miikkulainen, R., Mitri, S., Moriarty, D. E., Mouret, J-B., Nguyen, A., Ofria, C., Parizeau, M., Parsons, D., Pennock, R. T., Punch, W. F., Ray, T. S., Schoenauer, M., Shulte, E., Sims, K., Stanley, K. O., Taddei, F., Tarapore, D., Thibault, S., Weimer, W., Watson, R., Yosinski, J. (2018). The surprising creativity of digital evolution: a collection of anecdotes from the evolutionary computation and artificial life research communities. arXiv:1803.03453v2
- Langton CG. Computation at the edge of chaos: phase transitions and emergent computation. *Physica D: Nonlinear Phenomena*. 1990;42(1-3):12–37.
- Lenski, R. E., Ofria, C., Pennock, R. T., Adami, C. (2003) The evolutionary origin of complex features. *Nature*, 423:139–144.
- Ray, T. (1992) J'ai joué a Dieu et cree la vie dans mon ordinateur. *Le Temps strategique*, 47:68–81.
- Runco MA, Jaeger GJ. The standard definition of creativity. *Creativity Research Journal*. 2012;24(1):92–96.
- Sims, K. (1994) Evolving 3D morphology and behavior by competition. *Artificial life*, 1:353–372.
- Turing AM. On computable numbers, with an application to the Entscheidungsproblem. *Journal of Math*. 1936;58(345-363):5.