# The Role of Subjectivity in the Evaluation of Open-Endedness

Kenneth O. Stanley[1] and L. B. Soros[1]

[1]Department of Computer Science, University of Central Florida, Orlando, FL, USA
kstanley@cs.ucf.edu, lsoros@cs.ucf.edu

**Abstract**

Given the interest in artificial life in achieving open-ended dynamics it is important to establish what open-endedness means. Yet the definition of open-endedness has proven surprisingly difficult to pin down. This paper examines whether one reason for this difficulty may be that open-endedness is in part a subjective notion. That is, open-endedness may be in the eye of the beholder. Grappling with this question is ultimately important for the field so it can set its expectations realistically.

## Introduction

In the quest to achieve open-ended dynamics, a recurring question is how to measure the degree of success of any given attempt. This question is often central to the evaluation of artificial life (alife) worlds and their eventual publication; reviewers naturally seek reassurance that any claims to open-endedness are validated, pushing researchers to defend their chosen metrics. This demand for metrics in turn forces stakeholders to endorse *definitions* of open-endedness (Bedau et al., 1997; Channon, 2003; Juric, 1994; Maley, 1999) because the appropriate metric for evaluating any given phenomenon depends intrinsically upon the definition of that phenomenon.

For the field of open-ended evolution (OEE), achieving consensus on a definition has proven surprisingly contentious. Some have pursued a rigorous quantitative path, aiming for precision. For example, *activity statistics* (Bedau et al., 1997, 1998) conceive open-endedness as the persistence of innovative adaptive traits, which then facilitates the construction of a principled quantitative framework. Yet the results from activity statistics leave room for ambiguity, such that it remains unclear whether systems that pass their test actually deserve to be admitted as *truly* open-ended to the same degree as e.g. life on Earth (Channon, 2003). Others have taken a less equivocal path, relying on measures of certain kinds of change that are more effective at proving the *lack* of OEE than its definitive presence. Recently Dolson et al. (2015) proposed a set "barriers" to OEE, offering such a method for detecting the collapse of OEE but leaving aside how identify its presence with certainty.

The struggle to converge to a satisfying definition is perhaps surprising in part because OEE appears so strikingly dramatic compared to the more mundane processes of nature. The proliferation of life on Earth in seemingly inexhaustible variety and often increasing complexity over hundreds of millions of years presents a peerless phenomenon of prodigious creativity. So why is it so hard then to pin down formally what exactly OEE is?

One danger lurking behind the discussion of definitions is that a satisfying understanding of our perception of open-endedness might ultimately require grappling with a degree of subjectivity, an idea anathema to "science" in its conventional conception. Yet how can we really come to a consensus without allowing open discussion of some of the slippery subjective concepts that float just below the surface of discussions of OEE? For example, an interpretation of *interestingness* seems critical to a sincere recognition of OEE.

## Why Interestingness?

At issue here is not what the definition of OEE should be, but rather why it is so hard to agree upon *any* objective definition. A key problem with recognizing OEE or measuring its degree is that simply observing a succession of increasingly complex and novel artifacts, or even a succession of artifacts that pass a certain test like activity statistics, is not necessarily congruent with what we may want OEE to mean. For example, while the definition of complexity itself is subject to some disagreement and is in some cases uncomputable (Li and Vitányi, 2013), even if we accept that complexity increases in a system over some period of time under some acceptable definition, the question of whether the increases are *interesting* still looms (unless one equates complexity and interestingness, but that only begs the question again of what interestingness means). Importantly, this dilemma emerges for *any* objective measure of progress (even aside from complexity); for example, if we accept that a particular gene line adaptively expands its presence over some period, it is still presumably important whether that new line is in any sense interesting.

To see this issue more clearly, consider why OEE is wor-

thy of our scientific attention. For us to invest time in trying to understand and reproduce a key feature of natural evolution, we must at some level agree that it is interesting. But why is it interesting? What about what happened in nature makes it intrinsically interesting to us? While the grandeur of nature is aptly described through the pen of a poet or brush a painter, it is not clear that there is an *objective* answer to this question, which is problematic because presumably whatever the answer is, it is what helps to distinguish the most open-ended systems from those that are less so. After all, is it reasonable to classify a system that no human (including researchers in OEE) perceives as interesting as synonymous with what transpired in nature?

## In Pursuit of Ground Truth

Here is a thought experiment: Suppose we have a function for estimating the complexity of an artifact that we agree is sufficient for our purposes. Now imagine employing a random number generator to generate sequences of digits of arbitrary length. Starting at minimal complexity $K$, the generator continually outputs random sequences until hitting one (at random) that exceeds complexity $K$. Next, $K$ is incremented and the generator now generates new sequences until it hits upon complexity $K + 1$, and so on. In the end, the result is a succession of increasingly complex sequences of digits with respect to any given complexity measure.

While such a succession is increasingly complex, it symbolizes an intrinsically uninteresting system. Few would argue that such a system deserves a similar status as evolution on Earth. However, what if by extreme luck the generator happens to output a succession of sequences isomorphic to the sequences of DNA in a lineage of evolution on Earth, or to a set of Shakespearean sonnets? Does that then make the system more interesting?

The answer has to be no, even though these artifacts are isomorphic to the same artifacts that actually occurred on Earth. The problem is that these new artifacts are not generated in the right *context*. DNA *evolved* in the context of actual functioning life forms and Shakespeare wrote for an *audience*. Take away the context, and the system is nothing but monkeys getting lucky with keyboards – random luck.

Similarly, if through some sequence of unlikely events in some hypothetical universe a human body of precisely the same architecture as in our universe appears within a world that never had oxygen, while the complexity of the artifact is arguably equivalent to the complexity of humans in our universe, somehow the significance of the achievement is fundamentally lessened because the context is entirely inappropriate. OEE cannot be in this case the *explanation* for the interestingness.

But if humans (or Shakespearean sonnets) are not intrinsically interesting in all possible worlds, and sequences entirely isomorphic to sequences of artifacts that occurred in natural evolution can be fundamentally less interesting, then (1) *the artifacts produced by a system cannot be on their own the reason the system is interesting*. Furthermore, if a sequence of increasing complexity is not always intrinsically interesting, then (2) *increasing complexity itself cannot be the key determinant of interestingness either*.

These conclusions are consistent with the idea that *context* matters for deciding what is interesting. For example, a human body simply assembling into existence through a random process in an entirely inappropriate environment is outside of a context where it can be appreciated. In effect, if a process produces a human body within a context where that body can never actually function as a human being, the process overall is less interesting. And a sequence of increasing complexity is not on its own enough to be interesting. Thus it seems that, for particular artifacts to signify an interesting process, those artifacts must be the result of something that more genuinely expresses a high degree of OEE (which thereby provides a meaningful context for their emergence).

Yet here is the worry: This argument has become circular – it began with the idea the open-ended systems should be interesting in some way and now has ended with the idea that for a process and its products to be genuinely interesting it must be sufficiently open-ended. In other words, open-ended systems must be open-ended. Somewhere, the essential ground truth we are seeking seems to have slipped through the cracks of the argument. Where is the edifice upon which a non-circular definition can be built?

## Conclusion

The arguments herein do not prove that there is no satisfying objective definition, but they illuminate why it is so hard to settle on one. Moreover, they raise the unsettling possibility that there can never be such an objective definition because ultimately it may be that OEE and interestingness are inextricably intertwined such that ultimately OEE is as much in the mind of the beholder as interestingness. If we try to escape this subjective fate, we expose ourselves to the awkward proposition that systems dangerously close to random, or on the other end of the spectrum utterly facile, begin to pass our tests, defying all common sense and leading the field down a rabbit hole of self-congratulatory diversion. That is, ultimately we must be trying to achieve something interesting, or what is the point?

One potentially controversial solution could be to grapple with the subjective nature of interestingness and try to understand why we are so fascinated with evolution on Earth and OEE. The most interesting worlds are a potent combination of the artifacts within them and the story of how those artifacts interact with each other and the environment. Somewhere in that union is a narrative we find compelling – stories of survival and creativity – perhaps simply because these are also the stories of ourselves. While pursuing a phenomenon simply because we like it may seem dangerously

ungrounded, perhaps the very processes that produce things we like are those processes we actually need or want, and there may be nothing wrong with that.

## Acknowledgments

This work was supported by the National Science Foundation under grant no. IIS-1421925. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

Bedau, M. A., Snyder, E., and Packard, N. H. (1997). A comparison of evolutionary activity in artificial evolving systems and the biosphere. In *Proceedings of the Fourth European Conference on Artificial Life*, pages 125–134, Cambridge, MA. MIT Press.

Bedau, M. A., Snyder, E., and Packard, N. H. (1998). A classification of longterm evolutionary dynamics. In *Proc. of Artificial Life VI*, pages 189–198, Cambridge, MA. MIT Press.

Channon, A. (2003). Improving and still passing the ALife test: Component-normalised activity statistics classify evolution in geb as unbounded. In *Proc. of Artificial Life VIII*, pages 173–181, Cambridge, MA. MIT Press.

Dolson, E., Vostinar, A., and Ofria, C. (2015). What's holding artificial life back from open-ended evolution? *The Winnower*.

Juric, M. (1994). An anti-adaptationist approach to genetic algorithms. In *Proc. of First IEEE Conf. on Evolutionary Computation*, volume 2, pages 619–623. IEEE.

Li, M. and Vitányi, P. (2013). *An introduction to Kolmogorov complexity and its applications*. Springer Science & Business Media.

Maley, C. C. (1999). Four steps toward open-ended evolution. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-1999)*, pages 1336–1343.