

# Modelling and measuring open-endedness

Susan Stepney<sup>1,2</sup>

<sup>1</sup>York Cross-disciplinary Centre for Systems Analysis

<sup>2</sup>Department of Computer Science, University of York, YO10 5DD  
susan.stepney@york.ac.uk

## Abstract

Generating open-ended (OE) systems is a major and as yet unachieved goal of ALife research. Here I discuss aspects of defining, modelling, and measuring OE. I apply a simple model of OE to itself, thereby expanding the concept, to demonstrate how truly open and vast open-endedness is.

## Introduction

Defining, categorising and measuring open-endedness (OE) is a major goal of ALife research (Bedau et al., 2000; Packard et al., 2019a,b), typically in the context of evolution. In the introduction to the Artificial Life journal special issue on OE, Packard et al. (2019b) summarise recent work, including the York and Tokyo categories, which use concepts such as novelty, complexity, and ‘interestingness’ to capture the underlying intuitions of what OE is ‘really’. Indeed, the definition of OE itself often devolves into definitions of properties that are ‘intuitively’ properties of OE systems.

Here I discuss several issues arising from attempts to model and measure OE in general, not restricted to a purely evolutionary context. I describe a common 3-type approach to categorising OE, including possible routes for moving to higher types, providing examples of these kinds of categories from a range of disciplines. I next discuss the consequences of this model of OE both for simulations and for the definition of measures. I then apply this 3-type model to itself, to extend it and to describe a fuller OE model of OE. Finally, I suggest some routes forward, many of which are substantial research topics in their own right.

## Qualitative novelty

### Three types of novelty

Hochberg et al. (2017) define biological innovation as ‘a qualitatively new phenotypic trait that is associated with a step departure from an evolutionary trend’. It is important for OE definitions and measures to include this idea of *qualitative* change, a change in kind, not merely in number, where ‘more’ becomes ‘different’ (Anderson, 1972). Biology itself

distinguishes different kinds of change, such as variation, speciation, and major transitions (Maynard Smith and Szathmáry, 1995), the former one being mostly quantitative, the latter two being mostly qualitative changes.

One approach to defining OE is in terms of the production of a continual stream of novelty (Taylor et al., 2016); such a definition further requires a definition of *novelty* (Crutchfield, 1994). Novelty is something not seen before, or something unexpected from current trends; so it must be defined with respect to (a model of) what has been seen before, and what is expected.

Banzhaf et al. (2016) define novelty with respect both to a model (the concepts used to capture the form, structure, and behaviours of the entities in a particular system of instances) and meta-model (the concepts used to build the model). The approach is based on the notion that a sufficiently OE system eventually moves outside its model (or, more precisely, outside our model of it). They define three types of novelty: type-0 novelty, *variation*, as novelty within the model (such as new instances of a given type); type-1 novelty, *innovation*, as novelty that changes the model (such as speciation, new types of instances or behaviours); and type-2 novelty, *emergence*, as novelty that changes the meta-model (new concepts needed to define the model, such as a major transition). They require a system to exhibit at least **continual innovation** (not merely continual variation) to be considered OE, to capture the notion of *qualitative* change.

Taylor (2019) explores, expands, and modifies the Banzhaf et al. (2016) approach: he renames the novelty types as exploratory, expansive, and transformational; he defines novelty with respect to the original rather than current model; and he allows exploratory change to be included as a form of OE.

These three types of novelty correspond quite closely with Boden (1990, 2015)’s exploratory, combinatorial, and transformative types of creativity.

Figure 1 summarises the various names used for these concepts, including the ones used in the rest of this paper.

	type 0	type 1	type 2
biology	variation	speciation	major transition
Boden	exploratory	combinatorial	transformative
Banzhaf et al.	variation	innovation	emergence
Taylor	exploratory	expansive	transformational
<b>this paper</b>	<b>variation</b>	<b>innovation</b>	<b>transformation</b>

Figure 1: Various names used in the literature for three types of novelty

### Type-0: variation

Variation is exploration of a given model’s state space or landscape. This space may be discrete or continuous, finite or infinite. It might be bounded or unbounded. Boundedness implies some edge or barrier, some maximum beyond which one cannot go; unboundedness can move forever, sometimes returning to the starting point (as on the surface of a sphere), sometimes continuing to infinity (as walking along the number line). The space might be simple (such as a set of strings modelling a genome) or highly complex (such as a model of multiple complex phenotypes in a complex environment).

This concept of a boundary occurs across a wide range of disciplines. If there is an edge or boundary, there may be no other side, such as with a Moebius strip. The boundary may be an illusion, retreating as it is approached, as in the horizon. Or there may another side, but nothing there other than the void, chaos, wilderness, madness, the unknown and unknowable. The edge may be due to *structural or resource limitations*: size, number of components, energy, or other limits. The edge may be due to *perceptual limitations* of individuals, such as humans being unable to perceive unaided IR/UV light wavelengths or sub-sonic/ultrasonic frequencies. Different species of individuals may have different perceptual boundaries. The edge may be due to *cognitive limitations* of individuals, where the ‘other side’ is an incomprehensible domain.

Exploration within such spaces finds ‘more of the same’, variations on a theme. There are new instances, maybe of unbounded number, but they all fit the current model.

### Type-1: innovation

Innovation is a change in the state space or landscape that necessitates a change to the model of the system. Chipping away at an edge might lead to a breakthrough, demonstrating that it is instead a border between two domains, crossing which can open up a new state space for exploration. Such a border might be seen at the critical point of a phase transition, or at a bifurcation in a dynamical system.

A border might be created by bringing together disparate domains: the created ‘borderlands’ itself becoming a region ripe for innovation. The littoral region between land and sea, the twilight zone between night and day, an ecotone between ecological regions, a rite of passage between childhood and

adulthood, the crossroads, bridges and trade ports mediating contact between realms: these are all liminal regions of change, mixing, confusion, and flux, with potential for diversity and creativity.

The so-called ‘edge of chaos’ (Langton, 1990) is such a borderlands region, marking the border between the simple domains of order and chaos.

### Type-2: transformation

Transformation is a major change in the state space or landscape that necessitates a change to the meta-model of the system. The exploration has transitioned through a gateway into a new kind of domain, either discovered or created. This new domain cannot be modelled with existing concepts; the whole landscape is transformed in previously unrecognisable ways.

Literature has a whole subgenre of ‘portal fantasy’, with its secret gardens, its other worlds. Often in such literature if a person returns to the mundane world, they do so transformed by their experiences. Alternatively, literary portals may be one-way, and the traveller is fated to wander, potentially through further portals, with no chance to return home. Biology exhibits major transitions to higher level individuals, such as the transition to multicellularity, and other transformative gateway events, such as the Great Oxidation Event, when molecular oxygen became more abundant in the atmosphere, enabling new ways of making a living. Such events can transform individuals and environments, and are typically not reversible.

### Transitions between types

Type-0 variations are routinely explored in ALife simulations. Innovations, and moreso transformations, are much less common, and are more highly valued. However, these types may not be rigidly separated, but change may move from variation to innovation to transformation as it becomes more extreme.

Exploration can ‘push the boundaries’, or ‘push the envelope’. This may simply be exploring the space to the fullest extent, as with Novelty Search (Lehman and Stanley, 2008). Or it may be a way to make the current space larger, pushing the frontier, taming the wilderness and incorporating it into the known and knowable. It can also be used to move to a region that has a richer ‘adjacent possible’ (Kauffman, 2000), where innovation and transformation are more readily achievable.

A major source of borderlands innovation comes from ‘structure clash’ or ‘culture clash’ of incompatible models. The relatively simple models of each domain do not fit together; the borderlands model needed to reconcile them can be richly innovative. The domains should be sufficiently different that creativity is required to combine them, yet sufficiently similar that there is sufficient common ground for such reconciliation.

Borders also function to keep domains separate, limiting mixing and migration, allowing each to progress and change relatively independently. Borders prevent homogenisation and loss of diversity.

A borderlands region may start as a simple patchwork, but may become sufficiently complex to become a transformative gateway. Taylor (2019) suggests that ‘the distinction between an expansive [innovative] and a transformational novelty can be viewed as the difference between a door-opening novelty in the same domain and a door-opening novelty in a different domain’.

Packard et al. (2019b) describes major transitions as ‘characterized by an emergent hierarchy, with each new level in the hierarchy consisting of a new population of reproducing and evolving entities’ formed from lower level entities that originally existed and reproduced independently. Szathmáry (2015) breaks down a major transition into three phases: the formation, maintenance, and transformation of populations of higher-level individuals. He notes two types of major transition: fraternal, where the lower level individuals are alike (of the same type), and egalitarian, where they are of unlike (different types) and complementary.

Boden (2015) emphasises the importance of morphogenesis as a form of transformational creativity, with ‘the transformation of one already existing pattern into another, and then another, and yet another, and so on’, with the use of ‘complex, iterative feedback loops’. If the iterative developmental process is, or can become, self-referential, it opens the possibilities for computational reflection (Maes, 1987; Stepney and Hoverd, 2011; Hickenbotham and Stepney, 2016), where the model and meta-model can be actively accessed and manipulated through the code; and it can include ‘strange loops’, tangled hierarchies where the levels are not well-ordered (Hofstadter, 2007).

## Consequences for OE simulations

### Where the complexity resides

Rasmussen et al. (2001a) define dynamical hierarchies, where the higher order individuals have new and distinct properties not seen in their constituent parts. They note that, in a given simulation, higher and higher order systems and their emergent properties require ‘an appropriate increase of the object complexity of the primitives’. Rasmussen et al. (2001b) point out that this observation is in some sense trivial: ‘Assume that we have a minimal rule set that generates a particular dynamical hierarchy but only up to order  $N$ . If we stay within this simulation framework, it is necessary to add new rules to generate an additional order  $(N + 1)$  of emergence. How can the system generate a new, higher level of behavior unless something new is added to the elements?’. Each new major transition in biology, from replicating molecules up to societies, has been

based on the same underlying physics. Sub-symbolic Artificial Chemistries (Faulkner et al., 2018) have been developed as an approach to produce elements with a set of rich emergent properties.

Rasmussen et al. (2001b)’s need to add something new to the primitive elements is due to their desire for all levels of hierarchy to emerge purely from the elemental properties. Banzhaf et al. (2016) take a different approach, using not only a lowest *generative* layer, but also allowing simulation coding *shortcuts* to be used to capture higher level behaviours explicitly.

If individuals are to ‘push the boundaries’ of their own perceptual, behavioural, and cognitive limits, there needs to be something to make this worthwhile: something new to perceive, something newly achievable by new behaviours and thoughts. This implies the need for a great deal of complexity in the environment. Taylor (2019) also points out the crucial role of a complex environment, exploitable and modifiable by the individuals.

### Breaking the models

Although elemental complexity and shortcuts can be used in this way, the main challenge for simulation is the need to ‘break the model’ (Clark et al., 2005) in order to exhibit innovative novelty, and to break the meta-model to exhibit transformative novelty. Breaking the model potentially breaks any measure based on that model. There are two distinct types of models at work in Banzhaf et al. (2016)’s approach. First is the engineering model: a prescriptive specification of the simulation as built and executed (the specification of the genome and its expression, if you like). Second is the scientific model: a descriptive model of the observed outputs of the simulation (the model of the observed phenotypic structures and behaviours, if you like). What is observed, what is considered salient, is a choice; different observational choices, different perspectives, result in different models. Some models may include counterfactuals: ‘observations’ of entities or events that do *not* occur (Stepney and Hickenbotham, 2021).

OE novelties are defined with respect to the *scientific* model: what the simulation is emergently observed to do, not what the code is explicitly designed to do. So flocking can be observed as an emergent phenomenon (Reynolds, 1987), even though it is not explicitly encoded in the interaction rules.

So the scientific model can exhibit novel behaviours not explicitly encoded in the simulation. However, Crutchfield (1994) speaks of ‘intrinsic emergence’, where ‘there is a closure in which the patterns that emerge are important *within* the system’. In the context of simulation, this means that the observed novelties need to be captured in the *engineering* model. This allows them to be exploited by the simulation, for example through higher level selection pressures, for the generation of further higher level novelties. Banzhaf

et al. (2016) outline five stages to achieving full closure, of modifying the engineering model, from today's recognition outside the system with no change to the simulation, all the way to the emergent novelty being somehow emergently captured by the simulator itself. The latter stage is certainly technically challenging, and requires *self-modifying code*: new code to somehow incorporate and exploit the new novelties.

## Entities as code

This need for code change to capture emergent behaviours implies that the entities in the simulation should not be simple structures (as in typical agent based simulations), but should be the consequences of executing code. Here it is important to distinguish (at least) two levels of code (Hickinbotham et al., 2016): there is the 'physics' level, the underlying simulation engine, a form of interpreter animating the simulation entities; and there is the 'biology' level, the changing, varying, evolving individuals and environment, that can include code interpreted by the physics engine. The former code is fixed, and so should be minimised; the latter code can change, and needs to have the potential to be as computationally powerful as possible, up to Turing-complete.

Genetic Programming (GP) in its various forms (Banzhaf et al., 1998; Miller, 2011) evolves programs. These programs are typically the desired end point of an optimisation process: their fitness is evaluated by executing them. They could also be the active entities in an ALife simulation. GP can be used to evolve high level language expressions, or low level assembly language programs.

EvoMachina is an evolutionary system comprising a number of active code-based *machines* that perform operations like mutation, replication, and expression (Hoverd and Stepney, 2016). These machines are evolvable code entities.

Automata Chemistries are a form of Artificial Chemistry (Banzhaf and Yamamoto, 2015) where 'atoms' of assembly language instructions are strung into 'molecules' of assembly language programs. The program string is the genotype; the behaviour of the executing program is the phenotype. Examples include Tierra (Ray, 1992), Avida (Adami and Brown, 1994; Johnson and Wilke, 2004), Amoeba World (Pargellis, 2001; Greenbaum and Pargellis, 2016), and Stringmol (Hickinbotham et al., 2016, 2021; Clark et al., 2017). These systems typically exhibit a great degree of novelty, both innovative and transformative, as viewed in the scientific model and as evolved into novel code molecules in the engineering model.

Recent experiments with spatial Stringmol (Hickinbotham et al., 2021) exhibit increasing dynamical complexity as replicators evolve defences against parasites, parasites evolve mechanism to overcome these defences, replicators evolve defences against these new mechanisms, and so on. These processes demonstrate innovation in the engineering

model, and transformation in the scientific model (Stepney and Hickinbotham, 2020). Because the code is itself evolving, the resulting changed engineering model has to be abstracted from that code. The analyses to detect these novelties are currently performed outside the simulation itself (Stepney and Hickinbotham, 2021). Abstracting *minimal* such models is uncomputable if the models have Turing-complete computational power (for example, the minimal programs needed to specify Kolmogorov complexity are uncomputable).

In automata chemistries, the genome is the program, and the phenotype is the behaviour observed from the direct execution of that program. Another approach adding requisite complexity to an ALife simulation is to have a non-linear mapping from the genome to the phenotype; allowing the mapping to evolve (Taylor, 2019) adds another source of code-based complexity.

## Generating the models

**Explicit engineering.** Some authors engineer features into their systems to mimic an innovation or transformation. For example, Sayama (2019) engineers a *cardinality leap* in the size of the state space, from  $S$  to  $2^S$ , by explicitly facilitating formation of higher order aggregate entities. Moreno and Ofria (2019) describe their DISHTINY platform, which is designed to study fraternal major transitions by providing 'simple cell-like organisms with the ability and incentive to unite into new individuals'.

**Generative development.** The developmental stage from genome to phenotype provides a rich source of potential innovations and transformations, particularly if this mapping itself evolves (Taylor, 2019).

**Artificial Chemistries.** An Artificial Chemistry (AChem) explores its state space of possible 'molecules' in a generative manner: it has some underlying algorithm that generates new molecules from existing ones, through processes of composition and decomposition.

In the classical definition of an AChem (Dittrich et al., 2001), the form of this underlying algorithm is not broken down further. MetaChem (Rainford et al., 2020) provides a graph programming language approach to defining this algorithm. So we can think of an AChem defined using MetaChem as a graph model (algorithm) generating graph instances (molecules). We can then take this further, and have a meta-AChem, a graph meta-model generating graph models (AChems). See figure 2.



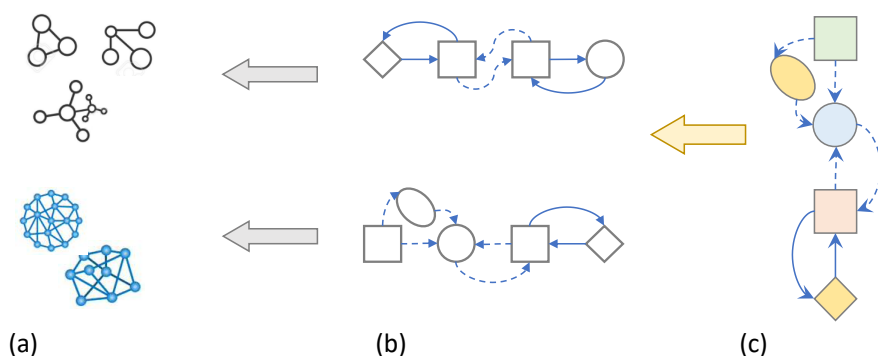


Figure 2: Conceptual figure: (a) families of molecules, represented as graphs; (b) the AChem algorithms that generate each family of molecules; an algorithm is represented as a graph in the MetaChem framework; (c) a higher order AChem algorithm, also represented as a graph, that generates the family of AChems as higher order molecules.

## Consequences for OE measures

### Quantitative change measures

Defining and using quantitative measures<sup>1</sup> to demonstrate the presence of OE is an active topic in the ALife community. In addition to classical measures such as Levenshtein edit distance, Shannon information, and Kolmogorov complexity, there are others hand-designed for ALife systems. Examples include those defined and used by [Bedau et al. \(1998\)](#); [Channon \(2006\)](#); [Droop and Hickinbotham \(2012\)](#). More recently, [Dolson et al. \(2019\)](#) have defined and implemented a suite of measures relevant to ALife systems.

All these aim to measure features associated with OE systems and processes, with the suggestion that a continual increase in the observed value of the measure implies an OE system. Few if any of these measures are explicitly linked to a particular model of OE; instead they are ‘intuitively’ thought to be linked to it.

[Hintze \(2019\)](#) critiques this use of continual increase of a measure, demonstrating it is not sufficient by measuring a specific trivial system, and showing it exhibits a continual increase in both diversity and complexity measures. Hintze’s system comprises a population of strings drawn from the alphabet  $\{R, L, F\}$ . The genome is the string; the phenotype is the 2D path drawn by interpreting the string as a LOGO-like program; the fitness is an explicit novelty measure that rewards a path that is different from those of other individuals in the population; mutation includes symbol change, deletion and insertion. He calculates a diversity measure, in terms of Levenshtein edit distance between

<sup>1</sup>Some authors use the term ‘metric’ to describe these. However, ‘metric’ has a specific mathematical definition, and many of the measures in the literature either do not demonstrate that they satisfy, or indeed fail to satisfy, that definition (particularly the triangle inequality condition). I use the term ‘measure’ instead. Although it too has a specific mathematical meaning, that is in a different domain, and so unlikely to result in confusion or over-interpretation of properties.

genomes, and a complexity measure, in terms of (a compression algorithm approximation of the uncomputable) Kolmogorov complexity of genomes. He finds that both these measures would indicate that this trivial system is OE.

Hintze makes a compelling case, and yet it can be critiqued in its own right. He points out that diversity can be measured at several levels – the alphabet, the genome, the phenotype – and is minimal at some levels and greater at others. He chooses to measure diversity and complexity at the genome (string) level, yet biology defines innovation at the phenotype level ([Hochberg et al., 2017](#)). (Of course, devising a measure of phenotype divergence or complexity is significantly more challenging than devising one for genome strings.) Kolmogorov complexity is uncomputable, which is why Hintze has to use an approximation; however, the approximation may well fail to capture exactly those ‘interesting’ cases that have low Kolmogorov complexity but high logical depth ([Bennett, 1998](#)). Measures should be computable, else approximations that do not have the theoretical properties of the stated measure will be needed. Furthermore, Kolmogorov complexity is a measure of randomness, not of ‘complexity’ in the complexity science sense. A measure like [Crutchfield \(1994\)](#)’s statistical complexity, which peaks at a value between pure order and pure chaos, would be more relevant.

Hintze concludes from his results that the criterion of an increasing measure is not sufficient for OE, and suggests we need better definitions of complexity and diversity that would reject his model. In terms of the 3-type model above, we can say that Hintze’s system exhibits unbounded *variation*, but no innovation or transformation. By [Banzhaf et al. \(2016\)](#)’s criterion, therefore, it does *not* exhibit OE; by [Taylor \(2019\)](#)’s relaxation of that rule, that includes continual variation, it does.

## Qualitative change measures

The main problem with the kinds of measures mentioned above is that they are measuring ‘Flatland’ (Abbott, 1884); they are confined *within* a model (even if that model is not always explicitly defined) and cannot see outside. They are measuring type-0 variational change. However, innovation and transformation *change the model*. We need measures that can detect when such changes are occurring, or are needed to adequately model the changing system. This is non-trivial; it is no coincidence that Dolson et al. (2019) state that they ‘would welcome a measurement of a system’s potential to produce major transitions in individuality’, that is, a measure of type-2 transformative change. But first, we should look for measures of type-1 *innovative* change.

Innovation may occur in the borderlands. These situations include ‘edge of chaos’ and other *phase transitions* (Badii and Politi, 1997, §6.3). Phase transitions have an associated *order parameter*, such as density or magnetisation in physical systems. The order parameter exhibits discontinuities (in value or derivative) at the critical point of phase transition. What are the relevant order parameters for innovative changes? Instead of looking for ever increasing measures, should we look for discontinuities in the measures? Simulation time does not look like a proxy for the independent variable, like temperature of magnetic field, here. Those variables exhibit one phase at low values and another phase at high values, with complexity at the critical value; simulation time moves from a relatively ordered phase to a critical point of innovation, where change occurs, and a new system results.

Another indication of model breakdown is divergence to infinity in a measure. Consider the example of the statistical complexity  $C$  of the logistic map, as examined by Crutchfield and Young (1990); Crutchfield (1994). The initial measure is the size of a stochastic finite automaton (SFA) needed to model an ensemble of strings of length  $L$ . Very regular strings, and also random strings, are modelled by small SFAs. Complex strings require larger SFAs, as they have more complex structure. The logistic map is shown to exhibit maximum complexity in the regions where it is on the border between periodic order and aperiodic chaos. At this point the complexity measure diverges with  $L$ : ensembles of longer and longer strings required larger and larger SFAs to model them: there is fractal-like ‘structure on all scales’. (Note this divergence is a function of system size, not of simulation running time.) At this point, the model (SFA) is no longer sufficient to capture the behaviour of the system, and a model with greater computational power, by including a stack, is needed. A *transformation* has occurred: the concept of stack has been added to the meta-model. As further transformations occur, the power of the model may need to be increased further, up to eventually Turing-complete.

## OE change measures

The measures described so far are all designed to work at the instance level, where variation is occurring. Changes to the model and meta-model occur when innovations and transformations occur, and in a fully OE system would occur autonomously. Measures of model change could therefore detect innovations; measures of meta-model change could detect transformations. Such measures would require that the models and meta-models be explicitly cast in some formal language, be comparable, and be measurable.

One potential approach is to cast models and meta-models in a graph-based language. This approach can support multi-level AChems (figure 2), is a flexible basis for evolutionary systems (Atkinson et al., 2020, 2021) and has a rich theoretical grounding in graph programming and graph rewriting.

Another potential approach is to examine the domain of Model Driven Architectures (Kleppe et al., 2003), and adopt some of their well-developed approaches to formal model transformation (typically expressed in UML-like languages).

## OE models of OE

Banzhaf et al. (2016)’s 3-type model summarised above provides an initial model with which to define OE. It is based on a landscape of possibilities (the state space, possibility space), and how that landscape can change. Variation is exploring that landscape, discovering the possibilities; innovation is modifying that landscape, finding new possibilities; transformation is forming a new *kind* of landscape, with new kinds of possibilities. Does it end there? If we think of this as one instance of a range of possible models of OE, we can try applying the concepts of this model to itself, to examine whether it can be used to expand the definition of OE.

**Variation.** Banzhaf et al. (2016)’s OE model defines three meta-levels: the instances (M0), the model (M1), and the meta-model (M2). The levels 0, 1, 2 could be considered as the beginning of *variation*, of the sequence 0, 1, 2, . . .; could we explore the number of meta-levels: could we have more meta-levels? In formal object oriented modelling, it is usual to have four levels: these M0–2 and a meta-meta-model (M3) (Kleppe et al., 2003, §8). It stops here, requiring all elements of M3 to be defined in terms of M3 concepts: the meta-levels have to top out somewhere to be implementable (and, admittedly, to be humanly comprehensible). But maybe we should allow a model of OE to explore multiple meta-levels, M4, M5, and so on.

If we have further meta-levels, we need to decide how to define further novelties. Innovation is a change to M1, induced by variation to M0. This change to M1 is a *variation* of M1. If this variation to M1 further induces a change to M2 (that is, is transformational), this is a *variation* of M2. So we can generalise. Let  $\text{var}_N$  be a variation of  $MN$  (so plain

variation is  $\text{var}_0$ ). Then an  $\text{innov}_N$  is a  $\text{var}_N$  that induces a  $\text{var}_{N+1}$  (plain innovation is  $\text{innov}_0$ , resulting in  $\text{var}_1$ ), and a  $\text{trans}_N$  is a  $\text{var}_N$  that induces a  $\text{var}_{N+2}$  (plain transformation is  $\text{trans}_0$ , resulting in  $\text{innov}_1$  and  $\text{var}_2$ ).

So we could extend Banzhaf et al. (2016)'s definition of OE to require 'continual production of each  $\text{var}_{0 < N}$ '.

We can also use this generalisation to help uncover and measure further innovations and transformations, as they have now been uniformly redefined as *higher level variations*: variations in the model and meta-model. We can use our understanding and measures of variation at M0 ( $\text{var}_0$ ) to define and measure  $\text{var}_1, \text{var}_2, \dots, \text{var}_N$ , and we can use our understanding of routes to generating  $\text{innov}_0$  and  $\text{trans}_0$  to understand routes to higher level innovations and transformations.

**Innovation and transformation.** This extension is one particular instance of a model of OE, built by varying the one described by Banzhaf et al. (2016). Can it be varied further? Can it be innovated? What is its meta-model? Its meta-meta-model? What further innovations and transformations of these models and meta-models are possible, and meaningful? These questions help illustrate how truly *open* and vast OE is (or should be).

**Change relative to what?** I noted above that Banzhaf et al. (2016) define change with respect to the *current* model, whereas Taylor (2019) defines it with respect to the *original* model. This gives rise to two very different concepts of OE. If change with respect to the current model is used, the very first glider seen in Conway's Game of Life (Gardner, 1970; Berlekamp et al., 1982) is surprising and novel, the model is updated, and every subsequent glider is merely another instance. On the other hand, if change with respect to the original model is used, every instance of a glider is as surprising and novel as the first. This may be an adequate approach for a system where each speciation event, each major transition, is hard won, and worthy of regard. But in a truly open system, multiple subsequent speciations are just 'more of the same', and not truly novel. Is a system that exhibits only more and more speciations (or even only more and more major transitions<sup>2</sup>) *truly* open-ended? The answer is a philosophical choice; I would say 'no', and so follow the Banzhaf et al. (2016) approach, as extended here.

<sup>2</sup>Banzhaf et al. (2016) present a meta-model (their fig.3) that treats composite entities at different levels as different concepts, and hence each major transition to a new level is a transformation. Subsequent major transitions *to the same level* are not considered as transformations, just innovations. That model is not 'minimal', as the concept of 'major transition' could be introduced once only, to cover all possible levels. However, that form was chosen specifically because of the central role of major transitions in biology. A different meta-model would not have to have that form.

## Conclusion

I have discussed some issues of modelling and measuring OE. These give rise to some suggestions for designing and analysing simulations for investigating OE. Such a system should at least include:

- a definition of the model of OE, with evidence it is open at all levels
- a definition of the measures in the context of that model, and a justification of how they capture OE properties
- measures of phenotypic structure and dynamics, not of genomes
- measures as a function of system size
- measures that indicate when the model needs to change/has changed, such as phase transitions and divergences
- measures of model and meta-model variation
- computable measures
- mechanisms to change the models (preferably intrinsically, but at least extrinsically) and associated measures
- individuals (and environments) whose measured properties depend on executing and changeable code

These suggestions merely scratch the surface of OE. There are many more research questions still to be resolved in modelling and measuring fully OE systems. We need radical new approaches to defining, generating, and studying them. These approaches should recognise that OE is an unbounded process, not some reachable end point.

**Acknowledgements.** I would like to thank the members of the York Complexity Reading Group – including Leo Caves, Jason Edwards, Philip Garnett, Divna Manalova, Tom McLeish, Ana Teixeira de Melo, Adam Nellis, Jeanne Nuechterlein, Elizabeth Tyler, and Richard Walsh – for many stimulating discussions of OE across a wide range of disciplines, some examples of which are given in this paper.

I would also like to thank my colleagues Wolfgang Banzhaf, Guillaume Beslon, Tim Hoverd, and Simon Hickinbotham for mind-expanding discussions over the past several years that have helped me refine some of my thoughts about meta-models, self-modifying code, and open-endedness in general. And further thanks to Simon Hickinbotham for perceptive comments on a previous draft of this paper.

## References

- Abbott, E. A. (1884). *Flatland: a romance of many dimensions*, by A. Square. Seeley & Co.
- Adami, C. and Brown, C. T. (1994). Evolutionary Learning in the 2D Artificial Life System "Avida". In *Artificial Life IV*, pages 377–381. MIT Press.
- Anderson, P. W. (1972). More is different. *Science*, 177(4047):393–396.
- Atkinson, T., Plump, D., and Stepney, S. (2020). Horizontal gene

- transfer for recombining graphs. *Genetic Programming and Evolvable Machines*, 21:321–347.
- Atkinson, T., Plump, D., and Stepney, S. (2021). Evolving graphs with semantic neutral drift. *Natural Computing*, 20(1):127–143.
- Badii, R. and Politi, A. (1997). *Complexity: hierarchical structures and scaling in physics*. Cambridge University Press.
- Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R., Foster, J. A., McMullin, B., de Melo, V. V., Miconi, T., Spector, L., Stepney, S., and White, R. (2016). Defining and simulating open-ended novelty: Requirements, guidelines, and challenges. *Theory in Biosciences*, 135(3):131–161.
- Banzhaf, W., Nordin, P., Keller, R. E., and Francone, F. D. (1998). *Genetic Programming, An Introduction*. Morgan Kaufmann.
- Banzhaf, W. and Yamamoto, L. (2015). *Artificial Chemistries*. MIT Press.
- Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C., Green, D. G., Ikegami, T., Kaneko, K., and Ray, T. S. (2000). Open problems in artificial life. *Artificial Life*, 6(4):363–376.
- Bedau, M. A., Snyder, E., and Packard, N. H. (1998). A classification of long-term evolutionary dynamics. In *ALife IV*, pages 228–237. MIT Press.
- Bennett, C. H. (1998). Logical Depth and Physical Complexity. In Herken, R., editor, *The Universal Turing Machine A Half-Century Survey*, pages 227–257. Oxford University Press.
- Berlekamp, E. R., Conway, J. H., and Guy, R. K. (1982). *Winning Ways for Your Mathematical Plays, Volume 2: games in particular*. Academic Press.
- Boden, M. A. (1990). *The Creative Mind: myths and mechanisms*. Weidenfeld & Nicolson.
- Boden, M. A. (2015). Creativity and ALife. *Artificial Life*, 21(3):354–365.
- Channon, A. (2006). Unbounded evolutionary dynamics in a system of agents that actively process and transform their environment. *Genetic Programming and Evolvable Machines*, 7(3):253–281.
- Clark, E. B., Hickinbotham, S. J., and Stepney, S. (2017). Semantic closure demonstrated by the evolution of a universal constructor architecture in an artificial chemistry. *Journal of the Royal Society Interface*, 14:20161033.
- Clark, J. A., Stepney, S., and Chivers, H. (2005). Breaking the model: finalisation and a taxonomy of security attacks. *RE-FINE 200, Guildford, UK; ENTCS*, 137(2):225–242.
- Crutchfield, J. P. (1994). The Calculi of Emergence: Computation, Dynamics, and Induction. *Physica D*, 75:11–54.
- Crutchfield, J. P. and Young, K. (1990). Computation at the Onset of Chaos. In Zurek, W. H., editor, *Complexity, Entropy and the Physics of Information*, volume VIII of *SFI Studies in the Science of Complexity*, pages 223–269. Addison Wesley.
- Dittrich, P., Ziegler, J., and Banzhaf, W. (2001). Artificial Chemistries—A Review. *Artificial Life*, 7(3):225–275.
- Dolson, E. L., Vostinar, A. E., Wiser, M. J., and Ofria, C. (2019). The MODES Toolbox: Measurements of Open-Ended Dynamics in Evolving Systems. *Artificial Life*, 25(1):50–73.
- Droop, A. and Hickinbotham, S. (2012). A quantitative measure of non-neutral evolutionary activity for systems that exhibit intrinsic fitness. In *Artificial Life XIII, East Lansing, MI, USA*, pages 45–52. MIT Press.
- Faulkner, P., Krastev, M., Sebald, A., and Stepney, S. (2018). Sub-Symbolic Artificial Chemistries. In Stepney, S. and Adamatzky, A., editors, *Inspired by Nature*, pages 287–322. Springer.
- Gardner, M. (1970). Mathematical games: The fantastic combinations of John Conway’s new solitaire game “life”. *Scientific American*, 223(4):120–123.
- Greenbaum, B. and Pargellis, A. (2016). Digital Replicators Emerge from a Self-Organizing Prebiotic World. In *ALife 2016: Cancún, Mexico*, pages 60–67. MIT Press.
- Hickinbotham, S., Clark, E., Nellis, A., Stepney, S., Clarke, T., and Young, P. (2016). Maximizing the adjacent possible in automata chemistries. *Artificial Life*, 22(1):49–75.
- Hickinbotham, S. and Stepney, S. (2016). Bio-reflective architectures for evolutionary innovation. In *ALife 2016, Cancun, Mexico*, pages 192–199. MIT Press.
- Hickinbotham, S. J., Stepney, S., and Hogeweg, P. (2021). Nothing in evolution makes sense except in the light of parasites. *bioRxiv*. <https://www.biorxiv.org/content/10.1101/2021.02.25.432891v1>.
- Hintze, A. (2019). Open-Endedness for the Sake of Open-Endedness. *Artificial Life*, 25(2):198–206.
- Hochberg, M. E., Marquet, P. A., Boyd, R., and Wagner, A. (2017). Innovation: an emerging focus from cells to societies. *Philosophical Transactions of the Royal Society B*, 372(1735).
- Hofstadter, D. R. (2007). *I am a Strange Loop*. Basic Books.
- Hoverd, T. and Stepney, S. (2016). EvoMachina: a novel evolutionary algorithm inspired by bacterial genome reorganisation. In *2nd EvoEvo Workshop, CCS 2016, Amsterdam, Netherlands*.
- Johnson, T. J. and Wilke, C. O. (2004). Evolution of resource competition between mutually dependent digital organisms. *Artificial Life*, 10(2):145–156.
- Kauffman, S. A. (2000). *Investigations*. Oxford University Press.
- Kleppe, A., Warmer, J., and Bast, W. (2003). *MDA Explained: the Model Driven Architecture: practice and promise*. Addison-Wesley.
- Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D. Nonlinear phenomena*, 42(1–3):12–37.
- Lehman, J. and Stanley, K. O. (2008). Exploiting open-endedness to solve problems through the search for novelty. In *ALife XI, Boston, MA, USA*, pages 329–336. MIT Press.
- Maes, P. (1987). Concepts and experiments in computational reflection. In *OOPSLA’87*, pages 147–155. ACM.
- Maynard Smith, J. and Szathmáry, E. (1995). *The Major Transitions in Evolution*. Oxford University Press.
- Miller, J. F., editor (2011). *Cartesian Genetic Programming*. Springer.
- Moreno, M. A. and Ofria, C. (2019). Toward Open-Ended Fraternal Transitions in Individuality. *Artificial Life*, 25(2):117–133.
- Packard, N., Bedau, M. A., Channon, A., Ikegami, T., Rasmussen, S., Stanley, K., and Taylor, T. (2019a). Open-Ended Evolution and open-endedness: Editorial introduction to the Open-Ended Evolution I special issue. *Artificial Life*, 25(1):1–3.



- Packard, N., Bedau, M. A., Channon, A., Ikegami, T., Rasmussen, S., Stanley, K. O., and Taylor, T. (2019b). An Overview of Open-Ended Evolution: Editorial Introduction to the Open-Ended Evolution II Special Issue. *Artificial Life*, 25(2):93–103.
- Pargellis, A. N. (2001). Digital life behavior in the Amoeba world. *Artificial Life*, 7(1):63–75.
- Rainford, P. F., Sebald, A., and Stepney, S. (2020). MetaChem: An Algebraic Framework for Artificial Chemistries. *Artificial Life*, 26(2):153–195.
- Rasmussen, S., Baas, N. A., Mayer, B., and Nilsson, M. (2001a). *Ansatz* for dynamical hierarchies. *Artificial Life*, 7(4):329–353.
- Rasmussen, S., Baas, N. A., Mayer, B., and Nilsson, M. (2001b). Defense of the *Ansatz* for dynamical hierarchies. *Artificial Life*, 7(4):367–373.
- Ray, T. S. (1992). An approach to the synthesis of life. In *Artificial Life II*, pages 371–408. Addison-Wesley.
- Reynolds, C. W. (1987). Flocks, Herds and Schools: A Distributed Behavioral Model. In *SIGGRAPH '87*, pages 25–34. ACM.
- Sayama, H. (2019). Cardinality Leap for Open-Ended Evolution: Theoretical Consideration and Demonstration by Hash Chemistry. *Artificial Life*, 25(2):104–116.
- Stepney, S. and Hickinbotham, S. (2020). Innovation, variation, and emergence in an automata chemistry. *Artificial Life 2020, Montreal, Canada (virtual)*, July 2020, 32:753–760.
- Stepney, S. and Hickinbotham, S. (2021). What is a parasite? defining reaction and network properties in an open ended automata chemistry. *Artificial Life 2021, Prague, Czech Republic (virtual)*, July 2021.
- Stepney, S. and Hoverd, T. (2011). Reflecting on open-ended evolution. In *ECAL 2011, Paris, France*, pages 781–788. MIT Press.
- Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences of the United States of America*, 112(33):10104–10111.
- Taylor, T. (2019). Evolutionary Innovations and Where to Find Them: Routes to Open-Ended Evolution in Natural and Artificial Systems. *Artificial Life*, 25(2):207–224.
- Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., Spector, L., Stanley, K. O., Stanton, A., Timperley, C., Vostinar, A., and Wiser, M. (2016). Open-Ended Evolution: Perspectives from the OEE Workshop in York. *Artificial Life*, 22(3):408–423.