

Social Systems and Ecosystems: History Matters

Inman Harvey

Evolutionary and Adaptive Systems Group, University of Sussex, Brighton, UK
inmanh@gmail.com

Abstract

In both social systems and ecosystems there is a need to resolve potential conflicts between the interests of individuals and the collective interest of the community. The collective interests need to survive the turbulent dynamics of social and ecological interactions. To see how different systems with different sets of interactions have varying degrees of robustness, we need to look at their different contingent histories. We analyse abstract Artificial Life models of such systems, and note that some prominent examples rely on explicitly a-historical frameworks; we point out where analyses that ignore a contingent historical context can be fatally flawed. Real life studies highlight the role of history, and Artificial Life studies should do likewise.

Introduction

In both ecosystems and social systems there are at least two levels at which, speaking loosely, ‘lifelike’ processes can be observed. There is one level at which the individual organisms, animals, humans are interacting with each other and pursuing their individual interests. But also there is a second level of ecosystem organisation, or social organisation, which provides the context within which they exist. In principle the same individuals could function, perhaps more or less successfully, if the ecosystem/social organisation was changed. One extreme version of such a change would be for the ecosystem/social organisation to break down in chaos, which is often against the interests of the individuals concerned. Systems survive or die, just as individuals do.

Ecosystems versus Social Systems. Social organisation can be the outcome of a social contract where individuals have chosen to agree to a set of rules. Ecosystems do not involve such explicit choice. Regardless of such differences, in both cases one can analyse individual behaviour in terms of self-interest potentially clashing with the interests of others around. In social systems we may call some actions ‘cheating’ and some consequences ‘punishment’. In ecosystems we tend to avoid such moral overtones and merely discuss ‘actions’ and ‘consequences’; the analyses may nevertheless be similar.

How are they Maintained? If a specific ecosystem/social system survives for a long time, explanation is called for. If no external authority is responsible for imposing this, then the organisation must be an emergent consequence of individual patterns of behaviour that are globally somewhat resilient to the perturbations of everyday life. We may ask how *one specific* ecosystem/social system is maintained, or we may ask about *generic* maintenance properties.

What is their Origin? Each specific ecosystem/social system will have its own unique history, from origins up to the present day; just as each organism has its unique genetic and

	Real Life	Artificial Life
Social systems	Bitcoin (Nakamoto, 2008) Common pool resources (Ostrom, 1990)	Iterated Prisoner’s Dilemma (Press and Dyson, 2012) (Stewart and Plotkin, 2013)
Natural systems	Ecosystems Niche construction (Clements, 1916) (Lewontin, 1969)	Daisyworld (Watson & Lovelock, 1983) (Harvey, 2015) Complex systems (May, 1972)

Table 1: Classes of decentralised social systems and natural (eco-)systems and their Alife counterparts analysed here.

developmental history. It is the main thesis of this paper that *generic* theories, that gloss over or average such *specific* histories, often fail to capture salient features of reality. Examples of such theories will be criticised.

Real Life. We consider both real systems and their artificial life counterparts, as in different columns of Table 1. Amongst real social systems we look at Bitcoin as a money transfer system, and common pool resource governance as studied by economists. Natural systems refer here to ecosystems as studied by ecologists in the field. It will be suggested that those studying such real life systems will have no problems agreeing with the thesis that history matters. Hence the main target audiences are those that produce abstract models that explicitly leave out any consideration of history.

Artificial Life Models. Simple abstract models of social systems are illustrated here by examples from IPD, Iterated Prisoner’s Dilemma. This is based on a classic two-person game where each player has simple choices and the interactions between them have consequences in terms of different payoffs. The basic dilemma of individual cheating versus cooperation is distilled into this simplest form. Ecosystem models discussed here include Daisyworld models where the organisms and environmental influences are characterised as variables interacting in a dynamical system.

Mathematical Summary. If processes are actually non-Markovian, modelling them as Markovian will lead to error.

Plain Language Summary. Real life takes place in a world of accumulated historical accidents that affect how social and ecological processes actually function. History matters.

Artificial Life Models of Governance

The Introduction to ‘Leviathan’ (Hobbes, 1651) gives the first known reference to artificial life under that name:

NATURE (the art whereby God hath made and governes the world) is by the art of man, as in many other things, so in this also imitated, that it can make an Artificial Animal. For seeing life is but a motion of Limbs, the beginning whereof is in some principall part within, why may we not say that all Automata (Engines that move themselves by springs and wheels as doth a watch) have an artificiall life?

This introduces the metaphor of a nation state as an artificial man, Leviathan, with different components functioning together as mechanically deterministic parts but forming a living whole. What sort of governance can provide some form of global harmony ensuring cooperation and collaboration between component parts and reconciling any potential conflicts between them? How does some form of social contract arise (and continue to survive) from a natural state of anarchy? Hobbes’ answer was for central rule by an absolute sovereign. Though such a sovereign is ultimately driven by his private interests, these are aligned with the public interests in so far as “The riches, power, and honour of a monarch arise only from the riches, strength, and reputation of his subjects”.

Here we follow Hobbes’ surprisingly modern notion of artificial life, and the use of models such as Automata in the study of real life governance. However we part company with him on his assumption of a need for a central sovereign. Leviathan is Hobbes’ exemplar of central authority, but in all further examples we discuss below there is no central authority, no rules for behaviour are imposed from outside.

Distributed Social Systems, Choice, Attractors

When governance arises solely through interactions between individual participants, different styles of governance can only arise through the different choices they make. Strategies any one individual has for choosing are typically conditional in part on the choices the others have made. These individual choices bind into a social system when there is a stable pattern that persists despite potential disturbances from within or without. In dynamical systems terms, we are looking for the attractors of such systems. There may be many possible such attractors, some more congenial than others to the participants – e.g. with higher payoffs in utilitarian terms.

Ecosystems, Choice, Attractors

With natural ecosystems we may not be considering the same type of explicit strategies of participants as in social systems. Nevertheless, a different type of choice is available, a choice of where to locate, which environment to inhabit. Animals may move from valley to hilltop; even plants, over generations, can shift their habitat. In this subtly different sense of choice, the component members of an ecosystem have ‘chosen’ to coexist in a specific locale where their various interactions (including their own knock-on effects on the environment) allow them to thrive. In the theoretical space of all conceivable ecosystems, there is a multitude of such viable and robust locales that act as potential attractors.

Real Social Systems

We focus here on two classes of real social systems with distributed governance: Bitcoin and common pool resource systems. Such systems are only stable if they are indeed currently near an attractor, which is another way of saying that they recover from small disturbances. We look at how such systems may adapt to changing circumstances over the longer term, and hence the role of historical contingency in how they come to be in one attractor rather than another.

Bitcoin

In commerce there needs to be common agreement about who paid what to whom; this is often centrally regulated by banks maintaining records. One version of this governance problem is that of verifying money transfers over the internet, and a very different style of solution is provided by Bitcoin (Nakamoto, 2008). Here the maintenance of book-keeping records is distributed, not centralised. The protocol used has to reconcile private interests with public interests; an individual would (dishonestly) benefit by spending the same money twice, but a money transfer system only works if such double-spending is prevented. Roughly speaking, this replaces trust in a sovereign central bank with trust in a majority consensus of multiple independent record-keepers distributed across the internet. This may be compared with a simple natural biological example where consensus amongst bacteria can be achieved via ‘quorum sensing’ (Miller and Bassler, 2001).

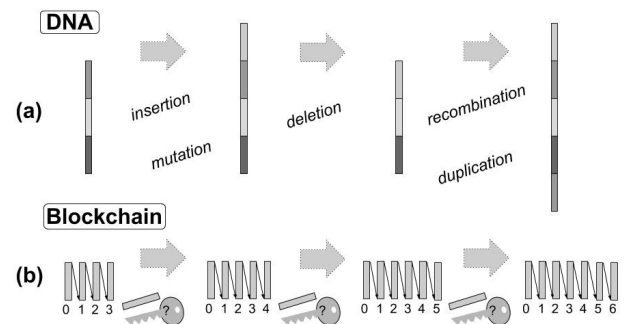


Figure 1: (a) DNA has statistical continuity over phylogenetic history, with noise. Older and newer data both matter. (b) Blockchain is built up systematically with new blocks added at the end, verified by consensus via ‘key-finding’ for each addition.

Blockchains and DNA. In Bitcoin the official record of all transactions in recorded history is maintained in a data object called a blockchain. Somewhat like DNA, this is a linear string of digits, meaning it is virtually free and instantaneous to copy. Multiple copies can be distributed widely. Like DNA, it can grow incrementally over time. Unlike DNA, the blockchain of accounting records cannot mutate or have parts excised; the protocol has to maintain accuracy and integrity across all copies of the blockchain as it is updated with new transactions (Figure 1). New transactions are bundled together

into a block to be added on to the end of the blockchain; then a deliberately lengthy and computationally expensive process is undertaken by each record-keeper to find a 'key' to validate it. This cryptographically-based key must identify both the old (mutually agreed) transaction history and the new block of transactions. Different record-keepers may have different updates to add (i.e. different new blocks), but the protocol must ensure agreement on just one of these as authoritative.

All record-keepers are competing to find a key for their version first, and since this task is deliberately designed to take several minutes, it will be clear who won the race. Though the key is hard to find, once found it is easy and near-instantaneous to distribute it with the blockchain for all other record-keepers to check it works. This verifies the integrity of past transactions and determines the official version of the new block of transactions. The system builds in incentives for the record-keepers to compensate for their efforts. Satoshi Nakamoto (2008), the pseudonymous inventor of the Bitcoin protocol, has invented a new form of social contract that survives (so far) through carefully defined incentives for the participants to put in the necessary work. The system is automatically policed against those that may want to subvert it. The potential of blockchains as distributed agreed records of history, where validation of updates can be agreed by consensus, goes far beyond this monetary Bitcoin example.

Common Pool Resources

Bitcoin has distributed, not centralised control. At a different scale, practical working examples of decentralised control can be seen in societies across the world ranging from water authorities in California to shared forest usage in Nepal and Switzerland and shared fisheries in Turkey. These are maintained and policed by the participants themselves rather than imposed by some external sovereign authority. Such 'common pool resources' have been the focus of economist Ostrom (1990). She proposes a list of design principles or 'best practices' that are common to such robust institutions:

- (1) Clear identified boundaries between those people and resources *within* the institution and those *outside*.
- (2) Appropriation rules congruent to local social and environmental conditions.
- (3) All (or most) members share in making or changing rules.
- (4) People who are users (or accountable to them) monitor the appropriation and resource management.
- (5) Sanctions for rule violations are graduated from low to high according to the severity or persistence of violations.
- (6) Conflict-resolution mechanisms are local and rapid.
- (7) External authority, e.g. higher government, does not enforce its own rule contrary to that of the local institution.
- (8) Where there are multiple levels of governance they are organised in multiple nested layers.

In such common-pool scenarios, unlike Bitcoin, anonymous entry or participation is not possible. The potential for the system to adapt itself according to changing local circumstances further differentiates this from what may be a serious weakness of Bitcoin. All participants have not only a stake in *maintaining* the rules (principles 4 and 6) but also in *changing* them (principle 3). Such adaptation in the governance system needs to be congruent with local social and environmental conditions (principle 2); and the social conditions may include further higher or lower level layers of

governance, overlapping in a nested fashion (principle 8). Within the generic constraints of these 8 principles there is scope for a multitude of possible governance systems each adapted, more or less, to local circumstances and fashioned through a historical succession of contingencies.

ALife Models of Social Systems: IPD

We move from stability, contingency, history in *real* systems to the same issues in ALife models. Recent innovations in IPD (Iterated Prisoner's Dilemma) provide a case study.

Motivation for IPD Models. These provide a minimal model of 2 agents ('prisoners') interacting. They must decide on actions independently, but the payoff to each depends on what they both decide, and is designed to provide a conflict between individual and collective gains.

The supposed story is that they have agreed beforehand to deny everything about some joint crime, but now they are interviewed separately by the police. Each has to decide whether to keep quiet as promised ('Cooperate' with the other prisoner) or make some deal with the police ('Defect'). In terms of utility, they both receive R (say 3) if both Cooperate; both receive P (1) if both Defect; and if one Defects, the other Cooperates the payout is T (5) to the defector and S (0) to the other. The choice of (T, R, P, S) = (5, 3, 1, 0) (Figure 2, following Press and Dyson, 2012) meets the PD condition $T > R > P > S$ that implies whatever agent 2's decision is, agent 1 would gain more by Defecting than Cooperating. The further condition $2R > T + S$ implies the total payout for both Cooperating, 2R, is higher than the total payment when one Cooperates and the other Defects.

The rules treat each agent symmetrically, so any difference in outcome depends solely on how their strategies interact. In a single game with no further consequences, each agent maximises their payoff by Defecting, irrespective of the other agent's choice. Hence they both Defect (D), receiving 1 each, whereas if both Cooperated (C) each would have received 3.

If such games are iterated indefinitely, in the IPD, then each agent's actions may influence future responses. Under some circumstances a regime of Cooperation for mutual benefit can arise; IPD studies usually focus on just what conditions allow this and discourage cheats (i.e. Defectors). Such conditions provide counter-examples to Hobbes' intuition that only a sovereign authority can guarantee a mutual Commonwealth.

Tit for Tat, TFT. A typical class of IPD strategy depends (either deterministically or probabilistically) on memory of the previous choices made by each agent in the previous N rounds, $N \geq 1$. Tit for Tat (TFT), for example, is the memory-1 strategy where an agent copies the action that the other agent took in the previous round (Axelrod, 1984; Figure 2b). Tit-for-Two-Tats is the memory-2 strategy where an agent only defects if the opponent defects twice in a row. More generally a memory-N strategy can be specified as a table with 4^N cells, relating to 4 possibilities CC, CD, DC, DD (for own +opponent choices) on N previous rounds, each cell specifying the probability that C will be chosen by this agent in the new round. For example, TFT with memory 1 has these probabilities of Cooperating, dependent on the previous round: CC 100%, CD 0%, DC 100%, DD 0%.

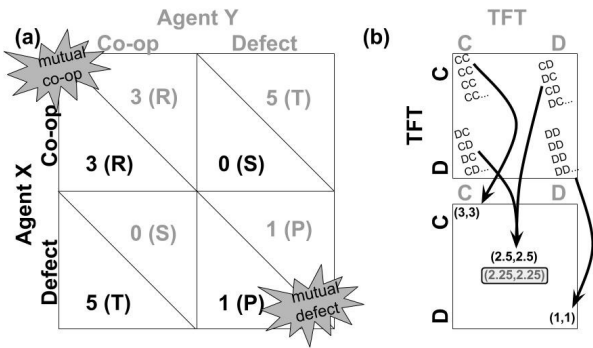


Figure 2: (a) The IPD payoff table (Press and Dyson, 2012). (b) Tit-for-Tat players differentiate into TFT_C or TFT_D (opening play C or D). Different varieties meeting (upper square) lead to 3 different end-attractors, average scores of (3,3), (1,1) and (2.5,2.5) (i.e. average of (5,0) and (0,5), lower square). The weighted (25%, 25%, 50%) average of all these attractor scores is different yet again, (2.25,2.25).

Historical and A-Historical Agents

Such memory-1 strategies depend explicitly on short-term memory of the previous round; but they also depend on long-term history of starting conditions, since the very first move makes a difference, say C for TFT_C or D for TFT_D . There are two possible routes to finesse this issue, the first being to acknowledge that TFT_C and TFT_D are indeed two different strategies with different consequences (Figure 2b). Only when TFT_C meets another TFT_C does the virtual circle of Cooperation take off. If both are instead TFT_D then a vicious circle of Defection takes over. A TFT_C meeting TFT_D results in alternating CD, DC choices. The starting conditions have a permanent effect on which basin of attraction is entered.

A second way to finesse this issue is to arrange affairs so that initial conditions eventually become irrelevant, and this could be the case with sufficient noise in the system. If with high enough probability a choice is accidentally reversed, then over enough iterations of IPD all possible basins of attraction will be visited. In a classic early Alife paper (Lindgren, 1991) explicitly used this method. There is a cost to be paid for finessing matters this way, however: TFT_C and TFT_D are now indistinguishable in such a theory, despite the fact that over any finite run they typically have totally different behaviours.

In principle the IPD game iterates for an arbitrary number of rounds, not known in advance. If both players know it to be the final game, this becomes a one-shot PD where both must rationally Defect. In turn, the penultimate game falls to the same analysis, and so on back to the first. An infinite series of rounds avoids this trap, but is impossible in practice. But we can have a finite, non-predetermined, number of rounds by arranging *after* each iteration a small (e.g. 1%) chance that it is *then* deemed to be the last. Then if any noise (as introduced by Lindgren) is small in comparison to this 1%, strategies such as TFT_C and TFT_D will be visibly seen to operate in different basins of attraction. Real world scenarios typically resemble this pattern rather than the infinite-iteration limit. For such real world scenarios, the history will matter.

This distinction between historical and a-historical agents is the central focus of this paper. Behaviour of the latter depends

only on recent short-term events held in ‘memory’, whereas the former also depends on one-off longterm origins in history. Crudely, this can be related to different perspectives from Biology and Physics. Typically many biologists are interested in a specific species or ecosystems with a specific evolutionary history (which we can relate to TFT_C or TFT_D). In contrast physicists, broadly speaking, may be happier making broad generalisations across some arbitrary range of entities (which we can relate to Lindgren’s TFT); often this makes the mathematics more tractable. Taken to extremes, this can result in broad statements that are generically true about “all possible organisms” assuming ergodicity, thus including extant organisms on this planet together with all extinct organisms, and indeed all conceivable organisms on all conceivable planets; but nevertheless misleading about any one specific non-ergodic organism. What is true about generic a-historical IPD agent TFT can be false about TFT_C or TFT_D .

Press and Dyson

A recent ground-breaking IPD paper (Press and Dyson, 2012), displayed a novel class of memory-1 ZD (Zero Determinant) strategies. These allow an agent — provided it no longer had the simple ambition to maximise its own payout that traditionally is expected in IPD — to tailor its strategy to guarantee that the opponent’s payout will average some value such as 1.5 (between P and R) regardless of how the opponent responds. Or such an extorting’ agent can guarantee that the excess of payoff above P will be shared in unequal proportions such as 3:1. The details of these ZD strategies are not discussed here. They are highly novel and counter-intuitive and are acknowledged by others to be valid, given the context; but many of the conclusions Press and Dyson drew have been shown to be misplaced (Stewart and Plotkin, 2013). We summarise these points, then go even further in questioning the validity of their Markovian assumptions..

Extortionate ZD Strategies. Suppose agent_X chooses an extortionate ZD strategy that gains a bigger proportion of the excess rewards (above a base-level of P) regardless of agent_Y’s responses. Then if agent_Y is an optimising player that adjusts strategy so as to increase its own payoff (Press and Dyson call this an evolutionary player) the result is that agent_X scores even higher. The erroneous implication Press and Dyson draw is that in an evolutionary scenario where multiple strategies are competing against each other, such extortionate strategies will triumph and dominate. As Stewart and Plotkin (2013) and other commentators point out, this is not so. If extortionate players came to dominate an evolutionary scenario, they will typically be competing with similar extortionate strategies. If agent_X and agent_Y are both forcing their excess payout (above $P=1$) to be 3 times greater than their opponents, this is neatly resolved by the excess being 0 for each, the (1,1) score of mutual Defection.

Generous ZD Strategies. It turns out that so-called Generous ZD strategies — that roughly speaking do the opposite of extortion in making sure that differential benefits mostly accrue to their opponents — will dominate in an evolutionary scenario. Such Generous strategies behave optimally against other Generous strategies, and also replace non-cooperative ZD strategies (Stewart and Plotkin, 2013).

Such ZD Strategies Ignore Historical Contingency

The main contribution of this paper to this novel development in IPD studies is to point out what other commentators have apparently missed: this whole class of ZD strategies, whether extortionate *or* generous, has been set up to be a-historical and hence to be largely irrelevant as models of human (or animal) strategies; since these are typically historical, contingent and contextual. Press and Dyson (2012) explicitly set up their ZD strategies to use the same finesse Lindgren (1991) uses, as discussed above, to average over all possible contingent long-term histories; they focus on generic strategies dependent on short-term memory alone. Indeed, they go further than Lindgren in showing that such Markovian assumptions allow any memory-N strategy to be generically equivalent to (some other) minimal memory-1 strategy.

Their proof covers the TFT strategy averaged over all possible histories, but fails cover a TFT_C strategy, even with its short history of a single first move. *A fortiori*, such IPD results have even less relevance to the real world when e.g. analysing the mating behaviour of *this* specific butterfly, with its long evolutionary and ecological history of multiple overlapping constraints as context; or when analysing the governance system for *that* Turkish communal fishing arrangement, with its long social and cultural history of multiple over-lapping polycentric social contracts. Ostrom (1990) explicitly mentions congruence with local social and environmental conditions among her design principles observed in long-lasting common pool governance systems, and this historical contingent contextually is what is stripped away in such generic mathematical proofs. Mathematically, one cannot analyse non-Markovian processes as if they were Markovian.

Real Ecosystems

We now consider the systems in the lower row of Table 1, starting with a minimal overview of real ecosystems.

Ecological Succession. This is the observed process of change in structure of an ecological community over the medium to long term. For instance after a mass extinction a typical sequence is for a few species of plants and animals to initially return; then successive new organisms arrive, building on what is already there in what Ostrom might want to call multiple nested polycentric layers in analogy to her social systems. In some cases this may be a somewhat predictable succession towards a final ‘climax community’ (Clements, 1916); but more recent ideas tend to take account of the many historical contingencies involved, including the varied feedbacks through knock-on environmental effects, and see a more unpredictable picture of ‘alternative stable states’ (Lewontin, 1969). In the short-term an ecosystem is in a stable steady state, but in the longer term it is somewhat accidental which one of many such possible equilibria it is, and what range of fellow organisms it contains.

Niche Construction. Such theories emphasise that organisms may not be merely accepting or selecting (through moving to) their specific environment; they may also have an active role in changing it (Laland and Sterelny, 2006).

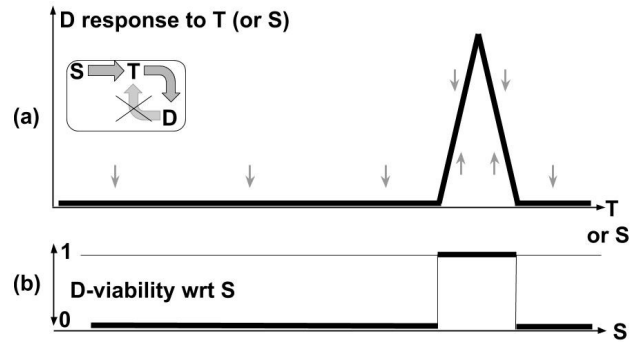


Figure 3: No-feedback scenario: environmental perturbation S (solar output) directly affects local env. T (temperature) which directly affects organism D (daisies). (a) D assumed to have steady-state dependency, ‘hat-shaped’ function of T, giving limited zone of viability. (b) D-viability (binary yes/no) plotted against perturbation S (here scaled to match T).

ALife Models of Ecosystems

Daisyworld (DW) models (Watson and Lovelock, 1983; Harvey, 2015) offer a simplified vision of how organisms and environment interact in some sense cooperatively. This can be compared to a very basic form of niche construction.

Motivation for Daisyworld Models. These are not widely known, and where known largely misunderstood (Harvey, 2015). The rationale is to model a number of types of organisms (e.g. one being ‘daisies’ D) that can survive within a limited range of local environmental conditions (e.g. one being ‘temperature’ T). Collective survival of an ecosystem of different organisms means *all* of them are currently viable in their local environment; *robustness* of an ecosystem is measured in terms of how wide a range of perturbing environmental conditions it can survive; e.g. an external ‘sun’ S creating hotter/colder conditions. An organism may have some local environmental effect (e.g. the albedo of a black daisy may raise local temperature), and *complexity* is measured as the number of such different effects within the ecosystem. The key DW result is that more such *complexity* leads to greater ecosystem *robustness*.

We demonstrate this, starting from the simplest ecosystem with a single species; underlying equations are in Appendix A. Fig.3 shows schematically the basic influence of environment T on an organism D. Figure 4 shows the consequence of further adding an effect from the organism D onto the environmental variable T. The consequence is to extend, i.e. widen the range of solar forcing (perturbing external effect S) for which the organism is viable (Harvey, 2015). Here the effect is positive (e.g. the albedo of a black daisy increases local temperature) the solar viability range is extended towards lower values than otherwise. A negative effect (e.g. white daisies tend to reflect heat and decrease temperature) would extend the solar viability range towards higher values.

Plus and Minus, Rein Control. Further, if both variants are potentially available with both positive *and* negative effects

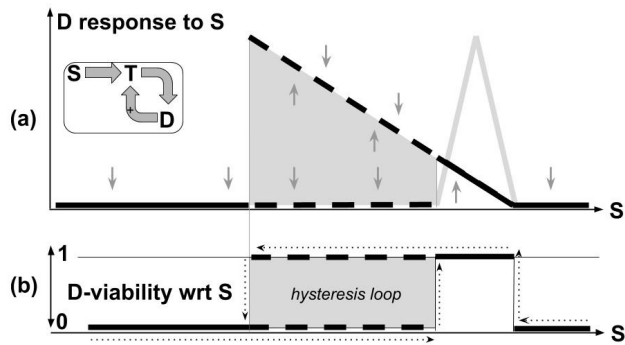


Figure 4: As Fig. 3 plus a further influence of D on T (here positive, black daisies increase temp.) (a) Peak response of D to changes in S is shifted, with a hysteresis loop. (b) D-viability zone is extended (here to the left) by a buffer zone, only effective if entered from high S values, and not low ones.

on the local environmental variable, temperature, they will collectively expand their joint eco-niche, as seen in Figure 5. This phenomenon depends on some basic assumptions spelt out in Appendix A; each variant, black or white, largely determines its own local temperature but with some 'leakage' between them in their shared environment. In this model, interactions between different 'species' such as D_B and D_W are only mediated via environmental variables, rather than through e.g. direct predation of one on the other. The results here, and developed further in Harvey (2015), demonstrate that any changes in viability range (for D_B & D_W , or D_B , D_W individually) always increase the range and never decrease it.

The expanded viability range takes the form of hysteresis loops as in Figure 5b. If the external perturbing force, here S, changes slowly, then which of the upper (viable) or lower (non-viable) arms of such loops is followed depends on which direction they are approached. In this sense history matters.

This is an example of 'rein control' (Clynes, 1969; Harvey, 2004). Clones observed a pattern when natural organisms exhibit homeostasis in response to external forces threatening viability both from above and below (e.g. both 'too hot' and 'too cold'). Rather than one mechanism responding in two directions, he noted two mechanisms each responding in one direction only. Since the reins of a horse have this same property, he called this 'rein control'.

This is further related to Le Chatelier's principle (Le Chatelier and Boudouard, 1898) as known to chemists and economists. This principle asserts that when any system in equilibrium is disturbed the system will adjust itself so as to (at least partially) nullify the effect of the change. A practical application of this principle is the use of a *buffer solution* which resists changes in pH when acid or alkali is added. These can be designed by chemists (Scorpio, 2000), or seen naturally where the *bicarbonate buffering system* regulates pH of blood in humans or other animals (Krieg et al., 2014).

Multidimensional Daisyworld

So far we only considered one environmental variable at a time: say temperature in DW, pH in the buffering examples.

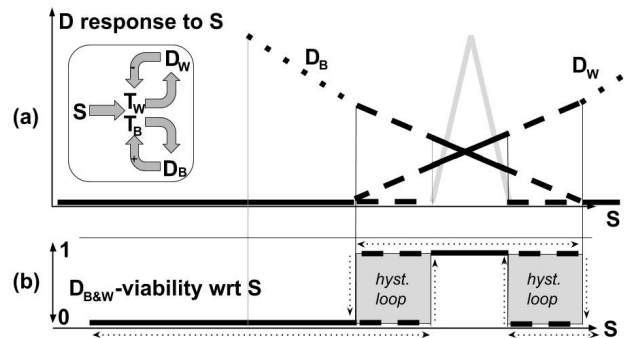


Figure 5: As Fig. 4 plus white daisies D_W affect negatively their local temp T_W as well as black daisies D_B affecting positively T_B . (a) shows steady-state values of each D (b) shows viability of D_B & D_W (simultaneously viable), against S.

What if two or more such variables are simultaneously relevant, e.g. both temperature and pH?

We can answer this within any very simple, abstract class of ecosystem models where (any number of) 'organisms' are modelled by 'hat-shaped' viability functions of (any number of) environmental variables; and in turn the organisms have any effect of any kind, positive or negative, on each or all of the environmental variables. In such cases it has been shown in the 'Gaian Regulation Theorem' (Harvey, 2015) that hysteresis loops or buffer zones as illustrated above exist regardless of the dimensionality of any such system. Perturbations in any number of dimensions will tend to be countered so as to widen — and never lessen — the viability range of any disparate group of organisms in an ecosystem, or of individuals in a corresponding social system.

As an abstract example, Figure 6a shows 8 groups of 8 species in clusters of narrow preferences for 3 environmental variables. In the absence of DW feedbacks at most one such group could be viable since the small viability spheres do not intersect (only P, V spheres shown here). If we add DW effects, different for all 8 members within each group, then when an external perturbation happens to pass the neighbourhood the whole group of 8 becomes jointly viable with a viability radius greatly expanded (from 0.05 to 0.218 for effect size 0.4; details in Appendix B). The expanded viability spheres (as V-sphere in Fig. 6b) may now overlap and (depending on environmental history) several such groups may become simultaneously viable. If the effect size were increased to 0.8, the mid-value perturbation C (0.5,0.5,0.5) would be within all 8 potentially expanded spheres, allowing all 64 (8x8) species with diverse environmental limits and diverse effects to be simultaneously viable.

This proof of principle still only has 3 dimensions of environmental variables, and is symmetrically set up to demonstrate an effect. Real systems will typically have more dimensions and be highly asymmetrical and locally varied, with convoluted overlaps of basins of attraction. Nevertheless we can see that different perturbation trajectories may result in very different ecosystems. Trajectories matter, history matters.

Such meandering paths through ecosystem-space can be compared with meandering evolutionary paths through DNA-

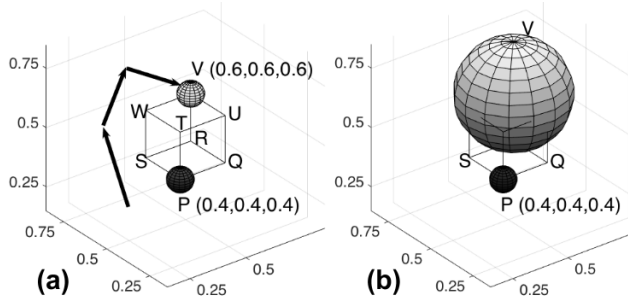


Figure 6. (a) 3 dimensions of external env. perturbations. P is group of 8 species, preferred env. (0.4,0.4,0.4), viable within radius 0.05 of this as shown by P sphere. Similar groups centred on Q, R...W, at corners of cube. Arrows show a possible trajectory of external perturbation. (b) This passes through viability zone of group V, DW effect consequently expands its viability radius to 0.218. See Appendix B.

space, and some degree of resemblance is not entirely accidental. From a high-level perspective the viability functions of DW can be related to the survival focus of Darwinian evolution. The natural settlement into attractors of the broad class of dynamical systems that is multidimensional DW relates directly to the natural selection of Darwinian evolution. Indeed the latter may be seen as a special case of the former. They both have surprising and counter-intuitive consequences.

Where an A-Historical Analysis Differs

The analysis of ecosystems in terms of DW, as presented here, is controversial (Harvey, 2015). One influential analysis (May, 1972) of an even broader class of ‘any large complex system’ (that includes multidimensional DW) purports to contradict it, proposing that, after some critical number of variables is exceeded, such systems are inherently unstable. Three mathematical flaws in this analysis have been previously exposed (Harvey, 2011). We here go further in identifying these flaws as arising from an a-historical analysis that resembles the a-historical analysis of IPD (Press and Dyson, 2012; Stewart and Plotkin, 2013).

May (1972) picks out an arbitrary equilibrium point of a large complex system and analyses its properties. This arbitrary choice, together with other explicit or implicit assumptions he makes, allows one to draw general conclusions; as the system gets larger, the chance that this specific equilibrium is stable tends towards the vanishingly small. This part of May’s argument resembles creationists’ arguments about the improbability of the ‘irreducibly complex’. But the fact that the probability of an arbitrary lottery ticket being a winner becomes arbitrarily small as the lottery itself gets arbitrarily big does not stop there being a winning ticket, or indeed many such.

A dynamical system left to its own devices will naturally head towards a stable equilibrium; any unstable equilibrium will only be briefly observed. As external conditions change, such a system inevitably passes through a sequence of metastable states. Thus any observed equilibrium is almost inevitably a stable one; which equilibrium it is depends on the history of the system. May’s analysis of a generic a-historical equilibrium state has little relevance for the analysis of

specific, observed, historically contingent equilibria (Harvey, 2011). Likewise the analysis by Press and Dyson (2012) of extortionate ZD strategies for IPD, or of Stewart and Plotkin (2013) of generous ZD strategies, has little relevance for historical contingent strategies such as TFT_C or TFT_D .

Conclusions

Crudely speaking, biology equals physics (and chemistry) plus history — stability in the short term plus the contingent context arising from an extended history of stability. More elegantly put, “Biology has always occupied a middle ground between the determinism of classical physics and the uncertainties of history” (Smith and Morowitz, 1982). When the physics of short-term stability is the focus of attention to the exclusion of contingent history, key concerns that can characterise complex systems can be missed.

It may be more than a coincidence that Press, Dyson, Stewart, Plotkin and May, variously cited and criticised above, all come from physics backgrounds. Another physicist, Rutherford (Birks, 1962), is quoted as saying “All science is either physics or stamp-collecting”. If the latter is interpreted as contingency, it need not be taken as derogatory; this is not only important for understanding real biology and social science but equally so for Artificial Life models of these.

In biological systems internal DNA is one obvious marker of a history, but other external markers may also be crucial. In polycentric social contracts (Ostrom, 1990) there may be multiple overlapping simultaneous systems of governance; likewise in polycentric organisms, polycentric ecosystems. Adaptations (and neutral changes) in any one system layer are within (and constrained by) the contingent current context of the others. Complexity of the whole arises through such adaptive/neutral trajectories through history, and cannot be explained a-historically.

A specific novel observation in this paper, apparently not noted by other commentators, is that the recently discovered extortionate ZD strategies in IPD (Press and Dyson, 2012), together with their generous cousin strategies (Stewart and Plotkin, 2013), have very little relevance to any biological or social studies of cooperation because they are all avowedly a-historical. Their Markovian assumptions are mathematically powerful but implausible as models of reality. The same applies to May’s (1972) analysis of large complex systems.

In passing we have noted that the blockchain of Bitcoin in its present form cleverly maintains the global history of transactions, and the full history is needed to establish the current state of accounts; in this sense the blockchain is historical. However the institutional framework of Bitcoin currently has no mechanism for adaptive change as per Ostrom’s principle 3; Bitcoin itself is a-historical.

Successful real social systems and ecosystems have a history of adapting to circumstances, and this gives context to their current stability. Artificial Life models should reflect this, and there are currently many promising research areas that give scope for developing many currently deficient a-historical models to take account of such contingency. History matters.

Appendix A

Fig. 5 shows ‘black’ and ‘white’ daisies, D_B and D_W , and respective local temperatures T_B , T_W (Harvey, 2004). Fig. 4, using D_B only, is similar except that D_W is clamped to 0.

Daisy viability w.r.t. local temperature is based on a ‘hat-shaped’ function $H(T)$ with (Fig. 3) peak value 1.0 at T_{opt} reducing to zero outside some limited viability range. Results are not qualitatively changed by different hat shapes.

$$(1) H(T) = \max(0, 1 - \text{abs}(T_{opt} - \alpha T))$$

Parameter α sets slope of hat. hence radius ($=1/\alpha$) of daisy-viability in terms of its local temperature. Parameter β sets the rate at which daisy-viability moves towards the hat-function:

$$(2) dD_B/dt = \beta (H(T_B) - D_B)$$

$$(3) dD_W/dt = \beta (H(T_W) - D_W)$$

The local temperature T_B , of black daisies D_B is based on the solar insolation S , altered (i) by *positive* influence from the black daisies, and (ii) by equilibration towards T_W . T_W is conversely affected, white daisies have *negative* effect. On the assumption that temperatures settle faster than rate of change of Daisies we can use the steady-state values as in (Harvey, 2004). Using T' for intermediate values of T , phase (i) is:

$$(4) T'_B = S + \gamma D_B$$

$$(5) T'_W = S - \gamma D_W$$

where γ parameterises the effect size for black/white daisies increasing/decreasing their own local temperatures. Phase (ii) gives the final temperature T as a compromise between each individual T' and their average current values; there is some ‘leakage’ (Harvey, 2004), here parameterised via δ (for $0 \leq \delta \leq 1$), between temperatures of black and white daisies:

$$(6) T_B = \delta T'_B + (1 - \delta)(T'_B + T'_W)$$

$$(7) T_W = \delta T'_W + (1 - \delta)(T'_B + T'_W)$$

If we choose $\delta=0.5$, then algebraic manipulation shows that equations (4,5) together with (6,7) can be replaced by:

$$(8) T_B = S + \varepsilon (3 D_B - D_W)$$

$$(9) T_W = S + \varepsilon (D_B - 3 D_W)$$

where for convenience we substitute $\varepsilon (= \gamma/4)$ for parameter γ .

Equations (1), (2,3) and (8,9) can be simulated computationally by choosing some specific value for S , and running these equations from starting values for D , T , until steady-state is reached. In hysteresis regions, the end-states reached will depend on the starting states. To plot one branch of each hysteresis loop, S should be initialised at a low value, and the computation run until D , T reach steady-state. Then S is incremented slightly, keeping *current* values of D , T as new starting values for the next run; this is further repeated, through to high values of S . If the process is then reversed, moving from high S to low S , the other branches of the hysteresis loops can be plotted. In Figure 5b, the viability of D_{B+W} is plotted as: IF ($D_B > 0$ AND $D_W > 0$) plot 1, ELSE plot 0.

Appendix B

Fig. 8 shows 3 dimensions of external env. perturbations. Viability of group of 8 species at P is 1.0 at (0.4,0.4,0.4), decreasing linearly to 0.0 at radius (Euclidean distance) 0.05. Each species has different +/- effects on 3 respective env. variables ($2^3 = 8$ variants); signs differ, but effect size is always 0.4. The other 7 groups (Q, ..., W) are formed similarly.

Effects of a P-species are multiplied by their viability and have two local contributions: half serves to shift the P-group

local env. away from the perturbing force (and is thus shared with other P-members; ‘leakage’); and half shifts the species-specific env. away from the P-local env. Over a trajectory of env. perturbations, at each point 20,000 computational iterations altered viability by 0.001 and local env. variables by 0.005 of their indicated shift. This smoothing of dynamics, together with the inheritance of previous env. values as perturbations changed, avoided numerical instabilities. A species was considered extinct if viability < 0.01 .

An effect size 0.4 expanded viability radius of each group from 0.05 to 0.218; effect size 0.8 expanded it further to 0.35.

References

- Axelrod, R., (1984). The evolution of cooperation. Basic Books, NY.
- Birks, J. B. (1962). Rutherford at Manchester. Heywood, London.
- Clements, F. E., (1916). Plant succession; an analysis of the development of vegetation. Carnegie Institute of Washington.
- Clynes, M., (1969). Cybernetic implications of rein control in perceptual and conceptual organisation. *Ann. NY Acad. Sci.* 156:629-670.
- Harvey, I., (2004). Homeostasis and rein control: from Daisyworld to active perception. In Pollack, J., Bedau, M., Husbands, P., Ikegami, T. and Watson, R. A. (Eds.), *Proc. 9th Int. Conf. on Sim. and Syn. of Living Systems, ALIFE 9*. MIT Press, Cambridge, MA.
- Harvey, I., (2011). Opening stable doors: complexity and stability in nonlinear systems. In Lenearts, T. et al., (Eds.), *Advances in Artificial Life, ECAL 2011*, pp 805-812, MIT Press.
- Harvey, I., (2015). The circular logic of Gaia: fragility and fallacies, regulation and proof. In Andrews, P., Caves, D., Dourest, R., Hickinbotham, S., Polack, F., Stepney, S., Taylor, T. and Timmis, J. (Eds.), *Proc. Eur. Conf. on Artificial Life 2015*, MIT Press.
- Hobbes, T., (1651). Leviathan. Andrew Crooks (publisher), at the Green Dragon in St. Pauls Church-yard, London.
- Krieg, B. J., Taghavi, S. M., Amidon, G. L., Amidon, G. E., (2014). *In vivo* predictive dissolution: transport analysis of the CO₂, Bicarbonate *in vivo* buffer system. *J.Pharm. Sc.* 103(11):3473-3490.
- Laland, K. N. and Sterelny, K., (2006). Perspective: seven reasons (not) to neglect niche construction. *Evolution*, 60(9), 1751-1762.
- Le Chatelier, H. and Boudouard, O., (1898). Limits of flammability of gaseous mixtures. *Bull. de la Soc. Chim. de France*, 19:483-488.
- Lewontin, R. C., (1969). The meaning of stability. *Brookhaven Symposia in Biology*, 22:13-23.
- Lindgren, K., (1991). Evolutionary phenomena in simple dynamics. In Farmer, J. D., Rasmussen, S. and Taylor, C., (Eds.), *Artificial Life II*. Edison-Wesley, Redwood City, CA.
- May, R. M., (1972). Will a large complex system be stable? *Nature* 238, 413-415.
- Miller, M. B. and Bassler, B. L., (2001). Quorum sensing in bacteria. *Annu. Rev. Microbiol.* 55:165-199.
- Nakamoto, S., (2008). Bitcoin, an electronic peer-to-peer cash system. Url: <https://bitcoin.org/bitcoin.pdf>
- Ostrom, E., (1990). Governing the Commons: the evolution of institutions for collective action. Cambridge University Press.
- Ostrom, E., Walker, J. and Gardner, R., (1992). Covenants with and without a sword: self-governance is possible. *American Political Science Review* 86(2), 404-417.
- Press, W. H. and Dyson, F. J. (2012). Iterated Prisoner’s Dilemma contains strategies that dominate any evolutionary opponent. *Proc. Nat. Acad. Sci.* 109(26), 10409-10413.
- Scorpio, R. (2000). Fundamentals of acids, bases, buffers and their application to biochemical systems. Kendall Hunt, Dubuque, IA.
- Smith, T.F. and Morowitz, H. J., (1982). Between history and physics. *J. Mol. Evol.* 18(4), 265-282.
- Stewart, A. J. and Plotkin, J. B., (2013). From extortion to generosity, evolution in the Iterated Prisoner’s Dilemma. *Proc. Nat. Acad. Sci.* 110(38), 15348-15353.
- Watson, A. J. and Lovelock, J. E., (1983). Biological homeostasis of the global environment: the parable of Daisyworld. *Tellus* 35B:284-289.