

Disentangling complexity from randomness and chaos

L. C. Zuchowski

June 8, 2011

Contents

1	Introduction	5
2	A genealogical definition of complexity	9
2.1	Complexity and statistical mechanics	9
2.1.1	Entropy measures	10
2.2	Complexity and computational sciences	12
2.2.1	Complexity, computers and structure formation	13
2.2.2	The institutional conception of complexity science	15
2.2.3	Emergent complexity	15
2.2.4	Measures of information and algorithmic content	18
2.3	Chaos theory and complexity	19
2.3.1	Chaos as complexity	20
2.3.2	Complexity at the edge of chaos	21
2.3.3	Statistical complexity measures	21
2.4	Disentangling complexity from randomness and chaos	24
2.4.1	A conceptual definition of complexity	24
2.4.2	From conceptual definition to complexity measure	26
3	Three models	27
3.1	A random system: The Kac ring	27
3.1.1	Dynamics of the Kac ring	27
3.1.2	Simulations of the Kac ring	30
3.2	A complex model: The Bak-Sneppen model	30
3.2.1	Dynamics of the Bak-Sneppen model	30
3.2.2	Simulations of the Bak-Sneppen model	32
3.3	A chaotic model: The logistic equation	34
3.3.1	The dynamics of the logistic equation	34
3.3.2	Simulations of the logistic equation	34
4	Comparing complexity: Shannon entropy and statistical complexity	37
4.1	Complexity estimates for the Kac ring	37
4.2	Complexity estimates for the Bak-Sneppen model	39
4.3	Complexity estimates for the logistic equation	41
4.4	Comparison of the models' complexity	45
5	Conclusions	47
A	MathCad codes for the simulations	49
B	Animations	53

Chapter 1

Introduction

During the last ten years complexity research has received a relatively large amount of attention by both the scientific community and the general public: paramount scientific figures like Nobel Laureate Murray Gell-Mann have championed its cause (e.g. Gell-Mann, 1994); Science magazine devoted a special issue to it; and in the 2004's bestseller lists *The Swarm* (Schaetzing, 2006) explored the sinister consequences of not paying enough attention to complex systems. One of the greatest draws of complexity as a field of research is the possibility of recognizing it in virtually every branch of science and the social sciences (e.g. Mitchell, 2009).

However, despite the labelling of an increasingly large number of models and natural systems as 'complex', the definition of the term has remained vague. Standish (2008) pointed out that these difficulties extend to both a qualitative identification of a conclusive set of defining features a complex system should possess, as well as to the quantitative measuring of complexity as a property. The difficulty of finding an unequivocal conceptualization and measure of complexity is also recognized by Gell-Mann (1995), who states that (p.1):

“[A] variety of different measures would be required to capture all our intuitive ideas about what is meant by complexity and by its opposite, simplicity.”

In fact, Lloyd (2001) enumerates forty-two different existing complexity measures in what is described as “a nonexhaustive list” (p. 7).

The lack of a well defined central concept appears to be closely related to the origin of complexity science as a conglomerate of spin-offs from a variety of scientific backgrounds. This is apparent in both the biographies of the field's main protagonists (e.g. Waldrop, 1992) as well as in the often repeated claim that complexity is present in a plethora of natural systems (e.g. Mitchell, 2009; Gregersen, 2003). Chu et al. (2003) suggested that the complexity researchers' strive for universality might preclude a concise definition of complexity and that one would have to settle for a qualitative description of com-

plex systems.

A closer review of the complexity measures compiled by Lloyd (2001), however, reveals that they borrow heavily from three particular other disciplines: statistical mechanics, computational sciences and chaos theory. Likewise, a closer look at the conception of these measures reveals that the struggle for a definition of complexity is largely identical with a failure to fully emancipate from these epistemological ancestors. A similar sentiment lies at the core of an extensive critique of the ‘complexity hype’ by Horgan (1995), who notes (p. 106):

“Complexologists have struggled to distinguish their field from a closely related pop-science movement, chaos.”

On a conceptual level, this implies that a successful definition of complexity will have to clearly differentiate a complex system from a stochastically random one, on the one hand, and from a deterministically chaotic one, on the other hand. A successful complexity measure should thus reliably distinguish a complex system from random and chaotic ones. In this thesis we will try to disentangle the notion of complexity from randomness and chaos. We will also examine the power of some existing complexity measures to distinguish a complex system from the other two.

Randomness and chaos Existing attempts at a quantitative definition of complexity rely on locating the concept between order and disorder. Thereby both chaotic and random systems are usually placed at the disordered end of the spectrum, while complexity is seen to reside somewhat closer to the ordered extreme (e.g., for review Ladyman et al., 2011; Edmonds, 1999).

A consequence of the identification of chaos and randomness with disorder is a widespread use of these three terms as synonyms in the canon of complexity literature. An example of this can be found in the subtitle of Waldrop (1992), which reads “The emerging science at the edge of order and chaos” - however, the intended contrast is clearly the one between order and disorder. Since we are explicitly trying to derive a definition of complexity by differentiating it from chaos and randomness, we will adhere to a more rigorous nomenclature. We will take the term *random* to describe behaviour caused by a probabilistic process that assigns equal probability to all possible outcomes. The descriptor *pseudo-random* will be applied to behaviour that resembles randomness but is not caused by a random process.

The term *chaotic* will exclusively be applied to systems that display pseudo-random behaviour and can be described by a deterministic, non-linear underlying set of equations. The definition of chaos itself is not a fully resolved epistemological problem yet. The current base-line formal definition is the one developed by Devaney (1989), which is summarized by Banks et al. (1992) in the following way (p. 332):

“Let X be a metric space. A continuous map $f: X \rightarrow X$ is said to be chaotic on X if

1. f is transient.
2. the periodic points are dense in X .
3. f has sensitive dependence on initial conditions.”

While conditions 1. to 3. in this definition are still debated (e.g., for review Smith, 1998), for our purposes it will be enough to note that ‘chaos’ is defined exclusively as a property of recurrent maps. As we will see in chapters 2 the dynamics of complex systems are not usually describable by such maps, in fact, they are not necessarily even deterministic. In chapter 3 we will find that chaotic mappings and complex systems are not directly comparable - our nominal attempt to do so requires the novel construction of an equivalent ‘chaotic system’.

We will therefore use the term ‘chaotic’ much more sparingly than is the norm in complexity science. In particular, we wish to avoid classifying any dynamics scheme that is ‘not linear’ (e.g. not additive) as ‘non-linear’ and therefore ‘chaotic’. Our definition of ‘pseudo-randomness’ is more inclusive and can be used for both deterministic as well as stochastic systems: chaotic mappings are a subclass of pseudo-random systems; a (cleverly) loaded coin would also be described as such.

In chapter 2 of this thesis we look at the roots of complexity in the three fields we identified as ancestor sciences: statistical mechanics, computational sciences and chaos theory. From this analysis we derive a qualitative definition of complexity. We also conduct a brief survey of the different complexity measures associated with each of the three areas. In chapter 3 we set up and explore computer simulations of three simple models: the random Kac ring model, the chaotic logistic equation and the complex Bak-Sneppen model of evolution. Chapter 4 compares the complexity values assigned to the three models by a traditional entropy and a statical complexity measure, respectively. From this we are able to draw conclusions about the appropriate use of quantitative complexity measures in chapter 5, which also hosts a general summary of our results.

Chapter 2

A genealogical definition of complexity

In the present chapter we trace the development of complexity conceptualisations in statistical mechanics (section 2.1), computational sciences (section 2.2) and chaos theory (section 2.3). Each section is divided into a qualitative part focused on distinguishing features of complex systems that the given field contributed and a quantitative part in which a sample of complexity measures is described. We fall short of discussing all forty-two measures accumulated by Lloyd (2001) but aim to include the most influential and representative ones.

2.1 Complexity and statistical mechanics

Statistical mechanics aims to relate the macroscopic behaviour of systems with a large number of microscopic constituents to the statistical behaviour of those constituents (e.g. Sethna, 2006). The pioneering work in this field was done by Boltzmann (1872, 1877a,b), who linked the statistics of gas molecules to their thermodynamic macroscopic gas properties.

The complexity encountered in systems studied in statistical mechanics seems to agree closely with the Oxford dictionary definition of complexity as “the state or quality of being intricate or complicated” (Oxford Dictionaries, 2011). Brownian motion (Brown, 1828), for example, appears intractable simply because there are so many particles interacting with each other that it becomes impossible for the human observer to account for them all. The interactions of the constituents in these systems appear random and undirected, i.e. the probability of interaction is same for any arbitrary chosen pair of particles.

The systems studied in statistical mechanics were also among the first in which scale-dependent levels of descriptions were discovered. Once translated via probabilistic measures and coarse graining to the

macroscopic level, the random interactions of the micro-constituents can be captured by macroscopic variables with much simpler time developments. In an early discussion of complexity in science, Weaver (1948) described the complexity of random systems in the following way (p. 539):

“From this illustration it is clear what is meant by a problem of disorganized complexity. It is a problem in which the number of variables is very large, and one in which each of the many variables has a behavior which is individually erratic, or perhaps totally unknown. However, in spite of this helter-skelter, or unknown, behavior of all the individual variables, the system as a whole possesses certain orderly and analyzable average properties.”

The notion of unexpected behaviour as a result of the interaction of a large number of constituents is the main inheritance that modern complexity research received from statistical mechanics. This includes the idea that the description of such a system is level-dependent. Furthermore, the physicists’ investigations into these systems also led to the first formal conceptualizations of order and disorder.

2.1.1 Entropy measures

The systems studied in in statistical mechanics are random or, as described by Weaver (1948), disorganized ones. The complexity measures developed for them therefore primarily aim at quantifying randomness and disorder. A hugely influential concept developed for this purpose is that of entropy, a thermodynamic variable that was first introduced by Clausius (1865) and defined in terms of statistical mechanics by Boltzmann (1877a,b).

Boltzmann’s first concept of entropy Boltzmann (1877a) derived the notion of entropy associated with a system of N particles confined to a volume V and with conserved energy E in an attempt to relate the second law of thermodynamics to the microscopic behaviour of the particles in an ideal gas. The phase space associated with this system of particles is $6N$ dimensional, assuming that each particle is described by a microstate vector $\mathbf{x}_n = (r_{1,n}, r_{2,n}, r_{3,n}, p_{1,n}, p_{2,n}, p_{3,n})$, where \mathbf{r}_n denotes the position vector and \mathbf{p}_n the momentum vector of particle n . An additional assumption of energy conservation in the system guarantees that the phase space region under consideration is finite.

In the first version of the entropy defined for this set-up, Boltzmann (1877a) argued that the distribution of the particles through phase space - and thus the complete state of the system - should be described by a distribution function $f(\mathbf{x}, t)$. Thereby $f(\mathbf{x}, t)d\mathbf{x}$ describes the fraction of particles with phase space vectors x_n lying within the infinitesimal phase space volume $(\mathbf{x}, \mathbf{x} + d\mathbf{x})$.

Considering the dynamics of the gas over time, Boltzmann (1877a) then defined an associated quantity:

$$H(f) = \int f(\mathbf{x}, t) \log[f(\mathbf{x}, t)]d\mathbf{x}. \quad (2.1)$$

In order to evaluate H , it is necessary to add further assumptions about the behaviour of f with time and thus about the dynamical properties of the system of particles under consideration. The major conceptual difficulty thereby is the estimation of the effect the interactions of particles will have on the phase-space distribution f . In an earlier work, Boltzmann (1872) developed the *Stosszahlansatz* as a means of quantifying the effect of collisions in an ideal gas: it assumes that the particles interact solely by elastic two-body collisions and that their velocities are previously uncorrelated. Under these assumptions $H(f)$ can be evaluated and is found to decrease monotonically with time. The H -theorem and the *Stosszahlansatz* are crucial in rendering Boltzmann's statistical mechanics descriptions irreversible with time. They sparked vigorous contemporary criticism and are still a focus of the irreversibility debate today (e.g. Uffink, 2008, 2007).

Using H , the first version of Boltzmann's entropy is defined as

$$S_{B,1}(f) = -kNH(f). \quad (2.2)$$

If the H -theorem is held true, then (2.2) will increase monotonically with time and thus fulfil the requirements of the second law of thermodynamics.

Boltzmann's second concept of entropy Boltzmann (1877b) presented a coarse-grained cousin of the entropy $S_{B,1}$ derived in the earlier paper. The infinitesimal phase space volume $(\mathbf{x}, \mathbf{x} + d\mathbf{x})$ is now replaced by an element of a rectangular partition $w = \{w_i : i = 1, \dots, l\}$, where each of the l elements is of sub-volume Δw . Energy conservation of the system guarantees a finite phase space and thus the existence of a finite number l of partition cells. On analogy to the distribution function f , the state of the system can then be given by the distribution of particles within the cells, $D_w = (n_1, n_2, \dots, n_l)$, where n_i is the number of particles in the i -th phase space cell. Assuming that all particles are indistinguishable, the number of different arrangements of particles throughout phase space compatible with a particular distribution D_w is given by the multinomial

$$G(D_w) = \frac{\left(\sum_{i=1}^l n_i\right)!}{n_1!n_2!\dots n_l!} = \frac{N!}{n_1!n_2!\dots n_l!}. \quad (2.3)$$

The corresponding combinatorial (or coarse grained) Boltzmann entropy is defined by

$$S_{B,2}(D_w) = k \log[G(D_w)]. \quad (2.4)$$

It can be shown that $S_{B,2}$ and $S_{B,1}$ are equivalent (e.g. Frigg and Werndl, 2011, section 4) and that, under the same dynamical assumptions outlined above, $S_{B,2}$ will also evolve towards a maximum state (e.g. Boltzmann, 1877b; Ehrenfest and Ehrenfest, 1959).

Boltzmann's second concept of entropy has been recognized as the more general one (Carnap, 1977) and it is therefore $S_{B,2}$ which has been exported as a general phase space distribution measure, with applicability to a much wider range of systems than the thermodynamic ones studied by Boltzmann (1872, 1877a,b).

Since (2.4) was defined to capture the tendency of a truly random system, it is maximum for a uniform distribution of the system through phase space (i.e. all N_i being equal). It will be shown in section 2.2 that this corresponds to a random probability distribution. Entropy has been adopted as the primary measure of order and disorder by the Brussels' school (e.g. Prigogine and Stengers, 1984; Prigogine, 1980), who investigated some of the systems that are customary classified as complex. However, the fact that it assigns maximum values to a perfectly disordered state is often cited as a reason to reject classical entropy as a complexity measure (e.g. Ladyman et al., 2011). Revised measures that rescale the thermodynamic potentials to distributions with maxima shifted away from the disordered extreme include the long-range *mutual information*, which is defined as "the amount by which the joint entropy of two parts of a body exceeds the sum of their individual entropies" (Bennett, 2003, p. 37), and the *thermodynamic depth* (Lloyd and Pagels, 1988).

2.2 Complexity and computational sciences

Waldrop (1992), who sets out to write the history of complexity science and its major protagonists, clearly construes the origins of complexity theory to lie within the field of computational sciences. In particular, he attributes the discovery of complex systems to computer modelling pioneers like Christopher Langton, Donald Holland, Norman Packard and Stuart Kauffman. While the Santa Fe Institute is the centre of the complexity universe for Waldrop (1992), and his history of the field is in large parts identical with an institutional history of the research centre, *Mathematica* founder Stephen Wolfram constitutes another influential figure in the field with a similar background (Wolfram, 2002).

The study through computer models of simplified representations of biological (e.g. McNaughton, 1989; Kauffmann, 1995), economic (e.g. Arthur, 1989), physical (e.g. Bennett, 2003) or just artificially designed systems (e.g. Langton, 1990; Wolfram, 2002) led to a new conception of order and disorder. Thermodynamics fundamentally assumes that all large systems eventually tend towards disorder and a uniform phase space distribution (section 2.1). However, computational iterations of systems with a large number of interacting constituents often show at least localized 'structures' or ordered 'patterns'. The term is thereby used to describe a structure that can be clearly distinguished from all three (randomness, pseudo-randomness and chaos) disordered behaviours. The dynamics of the systems modelled vary greatly and can be probabilistic or deterministic (as we will see in chapter 3, a single complex system can be modelled by a stochastic and a deterministic version); and linear or non-linear (Lady-

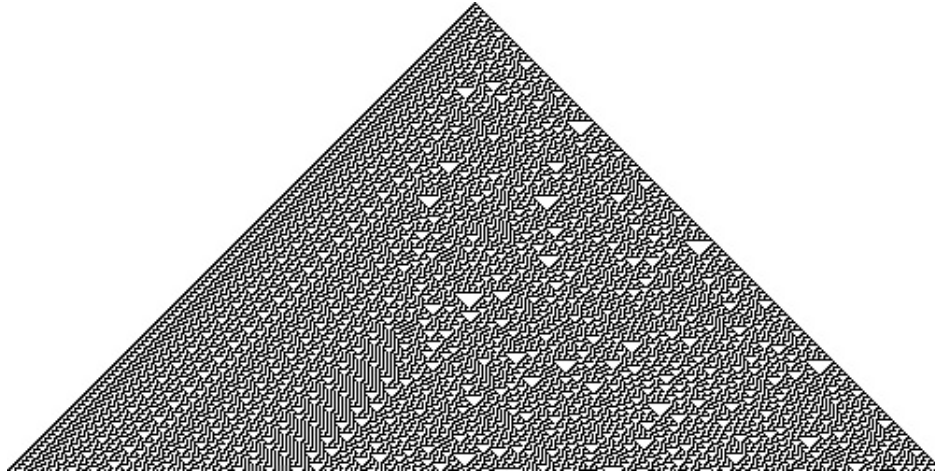


Figure 2.1: Wolfram (2002)’s rule 30 cellular automata phase-space diagram. Beginning from the top, consecutive iterations of the one-dimensional automaton are aligned. Image from Weisstein (2011).

man et al., 2011). Accordingly the behaviour that is commonly contrasted with structure formation is pseudo-randomness.

The discovery of these structures - made possible only by the introduction of sophisticated computer technology with graphical display powers (e.g. Coveney and Highfield, 1995) - added another layer to the definition of what would be considered a complex system. Scientists exploring these models began claiming the term ‘complex’ for systems displaying such surprising structure formation - thereby differentiating them from the mere stochastic systems studied by physicists like Boltzmann (1877a,b)

2.2.1 Complexity, computers and structure formation

It should be noted that complexity sciences’ staple terms, ‘organisation’, ‘order’ and ‘structure formation’ (Byrne, 2001), refer very specifically to patterns recognized in the models’ computer output. This often requires considerable restructuring of the data, so that the images shown for illustration might create the illusion of a spatial pattern in physical space, while in reality this order lies in abstract dimensions. For example, figure 2.1 - a diagram from Wolfram (2002) - plots the values of a one-dimensional cellular automaton versus time, creating a two-dimensional picture. Each row is thereby a copy of the automaton at a different iteration. The pattern we recognize is clearly the nested triangulation in the two-dimensional picture: in a physically correct one-dimensional representation of the row of cells that constitute the system these would manifest as fluctuations without an obvious geometrical counterpart.

In systems that are closely associated to nature and use similarly inspired metaphors for their patterns, the potential for confusion is even greater. Figure 2.2a shows a simulation of species’ co-evolution by Bak and Sneppen (1993). The dark patches, indicating extinction of one species and proliferation of another, are described as “avalanches” (p. 4085). However, the image again shows a linear arrangements of model elements plotted against time. By simple observation of the system alone, the variations

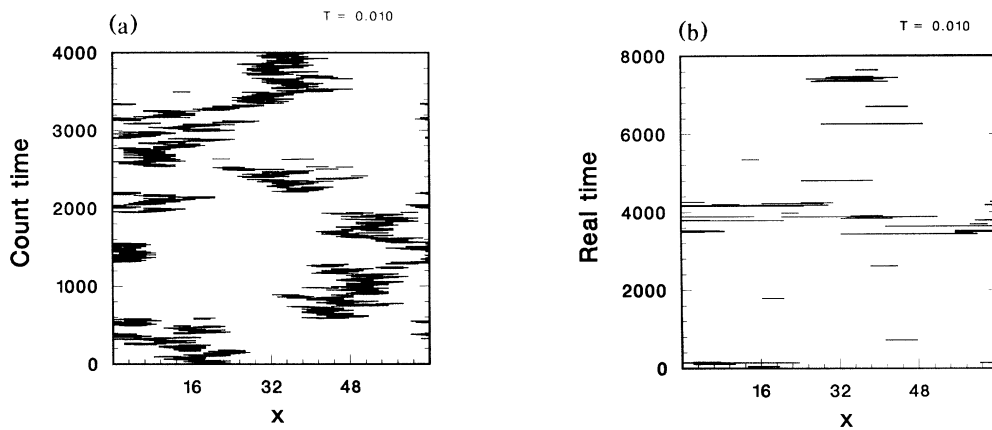


Figure 2.2: Phase-space diagrams of Bak-Sneppen avalanches. The one-dimensional system of elements x is plotted with a) time step and b) rescaled time. Image from Bak and Sneppen (1993).

would not resemble 'avalanches'. In fact, figure 2.2b shows that a simple rescaling of the time axis will greatly alter the observed pattern and render the descriptive metaphor unsuitable. The detection of order or structure in a complex system is therefore an abstract process of the detection of patterns in suitably manipulated model data - it is not necessarily linked to easily recognizable manifestations of the same classification in the actual model or the natural system it represents.

The importance of data analysis for the definition of complexity is recognized in an early discussion of the topic by Rosen (1977), p.29:

“We are going to define a complex system as one with which we can interact effectively in many different kinds of ways, each requiring a different mode of system description.”

In contrast, Edmonds (1999) proposed that an eventual definition of complexity would have to be based on the characteristics of the underlying model rather than an analysis of its output. James Cowan, first director of the Santa Fe Institute, expressed similar sceptical views when talking about the tendencies of researchers to overstress the similarities of their models to nature (quoted in Horgan (1995), p. 105):

“They say, ‘Look isn’t this reminiscent of a biological or physical phenomenon!’ They jump in right away as if it’s a decent model for the phenomenon, and usually of course it’s just got some accidental features that make it look like something.”

The new focus on digital pattern detection had a number of consequences: on the historico-epistemological level, complexity science as a field became virtually identified with the study of models displaying this structure generating behaviour; on a conceptual level, the focus on structure formation linked the notion of complexity to the well-established, philosophical concept of emergence. These two aspects will be discussed in section 2.2.2 and 2.2.3, respectively.

2.2.2 The institutional conception of complexity science

The institutional manifestation of the first, sociological, development is clearly the Santa Fe Institute, founded in 1985 by a group of scientists around former Los Alamos director James Cowan and Nobel Laureate Murray Gell-Mann (Waldrop, 1992). It was designed as a place where “questions could legitimately be asked about the emergent, adaptive, co-evolving behaviours many of these scientists observed in their own fields.” (Santa Fe Institute, 2011). In 2011 the institute’s research focuses range from “Physics of Complexity” to “Behavioural Dynamics”. However, while the choice of topics is interdisciplinary, all investigations carried out in Santa Fe have the common denominator of relying heavily on computer models. Using the choice of research technology rather than topic as the primary demarcation criterion is a distinctive feature of complexity science, albeit one that was possibly pre-empted by the ancestor field we will discuss in section 2.3, chaos theory (e.g. Kellert, 1993).

2.2.3 Emergent complexity

While the discovery of the structure-generating power of some representations of natural systems is the most celebrated achievement of complexity science, and constitutes a major part of the field’s “responses to the quest for cosmic meaning” (Davies, 2003, p. 13), we have discerned two distinct views of the nature of this emergence in the complexity community. Surprisingly, this divide seems to have found virtually no articulation in the vast amount of philosophically oriented literature on complexity.

Complex emergence I: From complexity to simplicity Researchers like Kauffman (1995), Strevens (2003) and (e.g. Prigogine and Stengers, 1984; Prigogine, 1980) stress the importance of emergent self-organisation in initially complicated, interconnected systems. These systems are defined as complex because “a vastly complicated assemblage of many small, interdependent parts somehow gives rise to simple large scale behavior” (Strevens, 2003, p. 1). Prime candidates for systems fitting this description are biological or economics one (e.g. Mitchell, 2009; Davies, 2003; Kauffman, 1995) and are often modelled as highly interconnected Boolean networks (Lansing, 2003).

Systems with complex-to-simple emergence are also often named *Complex Adaptive Systems (CAS)* (Holland, 1992; Lansing, 2003). The emergence in these systems is from a local to global level: many local components interacting with their local neighbours, according to potentially diverse rule sets, achieve self-organization to one simple collective action. Holland (1992) describes three mechanisms by which CASs create global emergent patterns (p. 26):

“Parallelism permits the system to use individual rules as building blocks, activating sets of rules to describe and act upon the changing situations. Competition allows the system to marshal its rules as the situation demands, providing flexibility and transfer of experience. [...] The procedures for adaptation - credit assignment and rule discovery - extract useful, repeatable events from this torrent, incorporating them as new building blocks.”

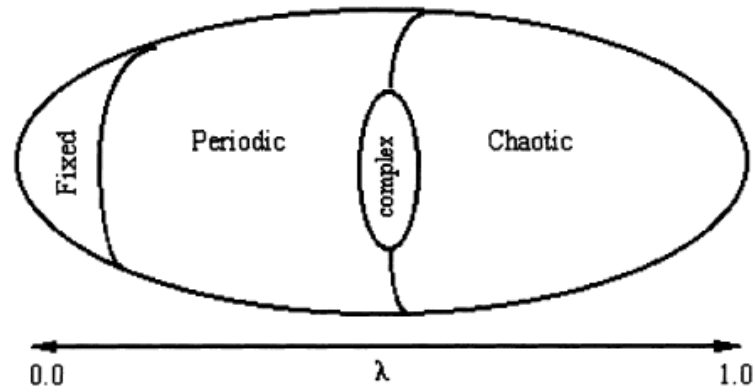


Figure 2.3: Order and disorder in the λ -parameter space. Image from Lansing (2003).

The concept of self-organisation to simplicity as described above is often also associated with the development of hierarchical structures (Mitchell, 2009, pp. 109-111). Another, less operational, mechanism for self-adaptation of a system is called *self-organized criticality*, which was discovered by Bak et al. (1989). It describes the behaviour of models that adjust themselves through episodes of short but intense change. The frequency of these ‘avalanches’ (section 2.2.1) scales with their intensity according to a power-law. As is often the case in complexity science, self-organized criticality models have been used to represent the dynamics of a wide range of processes (Frigg, 2003).

Complex-to-simple emergence is also the main feature of the second sort of complexity identified by Weaver (1948), which he contrasts with his ‘disorganized complexity’ discussed in section 2.1 (p. 5):

“They are all problems which involve dealing simultaneously with a sizable number of factors which are interrelated into an organic whole. They are all, in the language here proposed, problems of organized complexity.”

Complex emergence II: From simplicity to complexity The second school of complexity researchers roughly follows the belief expressed by Stephen Wolfram (quoted in Hotz (1997), p. 28)

“[V]ery simple rules, instead of producing very fairly simple behavior, actually produce extremely complicated behavior.”

The prime example for such systems are *cellular automata (CA)* (e.g., for review, Ganguly et al., 2003), a class of simple computer programs invented by Von Neumann and A. W. Burk (Ed.) (1963), which determine the development of values on a grid according to a set of interaction rules. John Conway’s Game of Life automaton has become widely popular and, despite the extreme simplicity of the rules that govern the virtual population, many dynamic and intricately complicated configurations have been discovered (e.g. Gardner, 1970). Wolfram (2002) conducted a systematic study of one-dimensional CA rules and discovered that there are many simple rule sets which produce complicated time developments. CAs can be categorized into four broad classes (Wolfram, 1984): Class I automata evolve to a homo-

geneous state; class-II automata develop periodic features; class III automata display pseudo-random behaviour; and class IV automata show complicated structures and pattern.

Langton (1990) introduced a concept called the λ -parameterisation of the CA rule space. Thereby rules which heavily favour one particular transition state (called the quiescent state) s_c are assigned low λ values, while high values of λ indicate that very few configurations of CA will lead to s_c . He found that class IV behaviour is displayed by CAs with rule sets characterised by medium λ values. For high λ values one obtains class III (pseudo-random) behaviour, for low ones the phase space outputs were ordered (class I) or periodic (class II). A further discovery of Langton (1990)'s scan through the parameterized rule space was the fact that the transition from ordered to disordered regime appears to be sudden rather than gradual. In analogy to the phenomenon in thermal physics, this has been named a "phase transition" (Langton, 1990, p. 24). The realisation that the complex CAs occupy only the small region of rule parameter space right before the order-disorder transition occurs, fits with the earlier hypothesis that the class IV set of rules should have zero measure (Packard and Wolfram, 1985). Figure 2.3 shows an illustration of this.

The parameter space location of the class IV automata also inspired the influential phrase of complexity being located at the "edge of chaos" (Langton, 1990, e.g. p. 36). However, since there is no straightforward way to identify the nearest-neighbours dynamics of CAs with the non-linear mappings we require for the label "chaotic" (e.g. Ganguly et al., 2003; Chua et al., 2002), at least under our strict requirements for nomenclature, the expression is a misnomer. More precisely, although not so linguistically refined, the class IV automata should be described as being located at the 'edge of pseudo-randomness'. We will explore in section 2.3 how the use of this expression has made it difficult to define complexity in relation to chaos.

Genetic algorithms CAs or other simple models with genetic algorithm appear to be a hybrid in our classification of emergent complexity. These are models whose rule sets are evolved towards a pre-set target performance over a series of runs, called generations (e.g. Holland, 1975; Goldberg and Holland, 1988; Mitchell et al., 1996). It is currently unclear whether genetic algorithm targeted towards a class IV end state tend to evolve towards rules with λ values close to the order-disorder transition (e.g. Packard, 1988; Mitchell et al., 1993).

We have classified the genetic algorithms as hybrids since the sub-programs they are evolving often have simple-to-complex emergence but the mechanism of genetic code development itself can lead to a subsequently more complex rule set, so that at the end of the generational cycle the system effectively has complex-to-simple emergence. For this reason, genetic algorithm models are primarily associated with complex adaptive systems (e.g. Kauffman, 1995; Mitchell, 2009).

However, while it is possible to match certain complex models with a certain type of emergence, the case of genetic algorithms also shows that phenomenological and dynamical simplicity is very much in eye of the beholder. The distinctions we have drawn are therefore schools of thoughts rather than confirmed ontological realities. As such, different researchers might describe emergence observed in the same model in different ways. This will be illustrated in section 3.2.2, when we will consider the emergent properties of the Bak-Sneppen model.

The two schools of complex-to-simple and simple-to-complex emergence in complexity science have also been identified with an anti-reductionist (e.g. McKelvey, 2001) and reductionist (e.g. Richardson and Cilliers, 2001) stance, respectively. Phelan (2001) pointed out that the complex-to-simple more than the simple-to-complex branch of complexity science invites accusations of constituting ‘pseudo-science’ (e.g. Shermer, 2002), mainly through an over-indulgence of the “resemblance thinking” (Phelan, 2001, p. 134), already criticised by James Cowan (quotation in section 2.1.1)

To sum up the qualitative part of this section: from the study of simplified computer models there arise two distinctions that differentiate complex systems from the random ones studied in statistical mechanics. (1) On the conceptual level, the models denoted complex are described by directed interactions. (2) Their phase space development is characterised by the existence of emergent properties. The emergent structures in (2) are not trivially predictable from the dynamics of the system - either because they show an autonomous degree of self-organisation (complex-to-simple emergence) or because their sophistication far exceeds that of the underlying set of rules (simple-to-complex emergence).

2.2.4 Measures of information and algorithmic content

In the later 1980s and early 1990s, the advent of complexity science coincides with a general “‘informatisation’ of biology” (Bawden, 2007, p. 315). Many of the complexity measures derived from the computer science heritage of complexity research are therefore information-based.

Shannon entropy Shannon (1948) developed a measure to calculate the amount of information gained from an ensemble of strings. In this framework, $m_1 \dots m_N$ is a set of N letters with a corresponding probability distribution $p_1 \dots p_N$ indicating the probability that the letter is sent. The discrete Shannon entropy is then defined as

$$S_{Sh} = - \sum_i p_i \log_2 p_i. \quad (2.5)$$

S_{Sh} fulfils the requirements for reasonable information measures (e.g Frigg and Werndl, 2011) and is maximal for a set of equally likely letters, i. e. $p_i = 1/n$. Several generalizations and extensions of (2.5)

have been developed (e.g. Beck, 2009).

If the probabilities p_i in (2.5) are interpreted as probabilities of finding a constituent of a given system in a cell w_i of a coarse grained phase space, then it can be shown that S_{Sh} and $S_{B,2}$ are equivalent (e.g Frigg and Werndl, 2011, section 4). This implies that both entropy measures of such a system will be maximal if a constituent is equally likely to be in any given phase space cell w_i . Rephrasing this in terms of information, it means that an outside observer has maximum uncertainty as to the phase space location of a given constituent (Jaynes, 1957a,b).

Algorithmic information content The concept of algorithmic information content of a specific binary sequence was developed independently by Solomonoff (1964), Kolmogorov (1965) and Chaitin (1966, 1969). In the context of complexity research, it is also often referred to as Kolmogorov entropy. The algorithmic information content S_K of a given binary sequence s is defined as the (bit) size of the minimum-length program s_U needed to compute s (e.g. Zurek, 1990):

$$S_K = |s_U|, \tag{2.6}$$

where $|x|$ indicates the bit size of a program x . In order to assure absolute minimality, S_K is defined for the Universal Turing Machine. However, as a consequence of the halting-problem discovered by Turing (1936), minimal programmes on the Universal Turing Machine cannot be determined and the algorithmic information content measures are fundamentally incomputable (e.g. Nannen, 2010).

Despite the fact that numerical values cannot be determined for (2.6), it is intuitively apparent that S_K will be maximal for a long random or pseudo-random sequence and minimal for a uniform one. As in the case of $S_{B,2}$ and S_{Sh} it has been argued that S_K is therefore not suitable as a complexity measure (e.g. Bennett, 2003).

A closely related measure that assigns minimum values to both extremes of order and disorder is the *logical depth* which is defined as the run-time of s_U on the Universal Turing Machine (Bennett, 1986). The *effective complexity* of a string s was introduced by Gell-Mann and Lloyd (2004) and denotes the algorithmic information content of just the regular part of s .

2.3 Chaos theory and complexity

As was the case with computational sciences (section 2.1), the roots of complexity research in chaos theory are biographical as much as epistemological. A sizeable contingent of the early complexity community, including Doyné Farmer, James Crutchfield and Norman Packard, recruited itself from the “Dynamical Systems Collective” (Gleick, 1987, e.g. p. 241), a group of young physicists at the

University of California who had been amongst the first to systematically explore and describe non-linear equations. While chaotic systems of equations produces the similar pseudo-random output as some complex models (section 2.2), they possess both a clearly defined mathematical structure (see chapter 1 for a formal definition) and a consistent set of mathematical tools for their description (e.g. Kellert, 1993; Smith, 1998; Hilborn, 2002). However, as Gregersen (2003) puts it, “both fields are spinoffs of the new revolutionary use of computers in science” (p. 207), which made the chaos researchers naturally allies (or even automatic members) of the complexity community.

2.3.1 Chaos as complexity

Many complexity researchers seem to naturally include chaotic systems under the umbrella of complexity. For example, Davies (2003) states that “among the earliest complex systems to be studied were those we now refer to as chaotic” (p. 25). Including chaos in the definition of complexity creates a number of technical difficulties, which are identified by Gregersen (2003) in the following way (p. 208):

“Whereas the distinctive trajectories of chaotic systems are highly contingent upon the exact values of the initial conditions, self-organized complex systems are more robust, that is, they can take off from a broad variety of initial conditions. Moreover, ordered complexity only emerges in rather limited domains of the chaotic systems governed by the Lyapunov exponent.”

It is not clear from the above quote, or indeed the chapter it is taken from, what is meant by “ordered complexity”. However, the expression seems to hint at the issue that we have identified as the major obstacle towards including chaos in a definition of complexity: the fact that, even allowing some liberties in the choice of an appropriate phase space and the underlying dynamics, the emergent property of a chaotic system is pseudo-randomness rather than structure (e.g. Hilborn, 2002; Newman, 1996). In fact, Werndl (2009) has shown that an observationally equivalent stochastic process can be found to simulate the output of any chaotic system. From the discussion in section 2.2 it is clear, however, that the construction of complexity as a science and as a concept independent from previous attempts at studying large systems, hinges crucially on the existence of phase-space portraits that distinguish themselves from mere pseudo-randomness. Including chaos in a complexity definition thus either implies greatly weakening the distinction between statistical mechanical and complex systems or making the awkward proclamation that ‘non-linear, deterministic pseudo-randomness counts as complex but other pseudo-randomness does not’.

A further difference between chaotic systems and those defined as complex in section 2.1 is the fact that chaotic maps as defined in chapter 1 lack the mechanisms of inter-connectedness and self-organisation to produce complex-to-simple or simple-to-complex emergence. Since non-linearity in these maps arises from self-feedback, chaos is generally not related to the existence of a large number of constituents (a

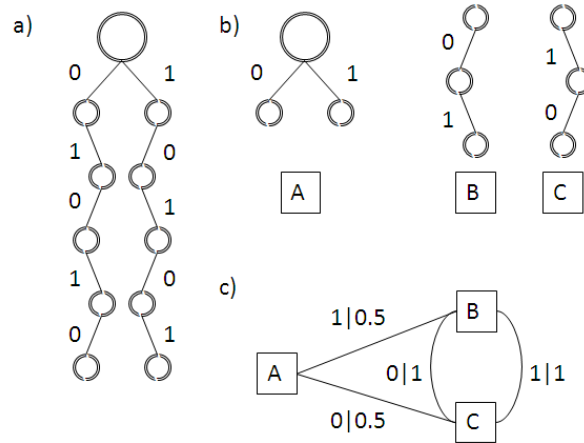


Figure 2.4: Computing the statistical complexity of the string 010101... a) Parse tree. b) Morphs. c) Reconstructed parse tree with probabilities indicated above the edges.

smattering of prominent examples illustrating this: the chaotic pendulum, the Smale horseshoe map, the Hennon map). Here again a crucial mark of complexity as derived from computer sciences and statistical mechanics would have to be relinquished in order to accommodate chaos.

2.3.2 Complexity at the edge of chaos

As described in section 2.2.3, we consider the expression “the edge of chaos” to be a misnomer and would prefer a rephrasing as ‘the edge of pseudo-randomness’. At the very least this would clear up the internal logical inconsistency that chaos in some contexts is described as complex (e.g. Davies, 2003; Mitchell, 2009) while at the same time Kauffman (1995) contains the following illustration of the relation between ‘chaos’ and complexity (p. 90):

“But at the edge of chaos, the twinkling unfrozen islands are in tendrils of contact. [...] Yet since the system is at the edge of chaos, but not actually chaotic, the system will not veer into uncoordinated twitching.”

While the term ‘chaos’ clearly fits better with the poetic language of the passage, there is no reason to assume that the potential processes causing ‘uncoordinated twitching’ of light bulbs in a Boolean network are non-linear (or in the general case, deterministic). ‘Pseudo-random’ would clearly be a better description and also avoid having to occasionally describe chaos as complex while at the same time placing complexity at the edge of chaos.

2.3.3 Statistical complexity measures

While our conceptual analysis above clearly suggests that complexity science should disentangle itself from the field of chaos research and ditch the unsuitable metaphor of “the edge of chaos”, research on complexity measures inspired by chaos theory has yielded the only alternative strand to the traditional entropy measures oriented on randomness.

Crutchfield’s statistical complexity The concept of statistical complexity was developed by Crutchfield and Young (1989) in an attempt to integrate the notion of structure emergence (section 2.2.3) as a crucial feature of complexity into a quantitative measure. The aim is thereby to “adequately capture the correlational structure in its behaviour” (Feldman and Crutchfield (1998), p.1), while keeping in mind that “both maximally random and perfectly ordered systems possess no structure” and should thus be assigned vanishing or low complexity values. The general principle behind statistical complexity is the compilation of a statistic of sub-strings (or patterns) in a the suitably reorganized output of a model system. A rough outline of the algorithm is the following:

- The system is decoded into a string with an alphabet of N symbols s_k . In most cases this means coarse graining the model’s phase space into N cells; e.g. for a system with the output of a simple sine curve a partition of $P = \{(0, 0.5]; (0.5, 1)\}$ could be assigned. If an element of the system is in the lower cell it will be represented by 0, if it is in the upper cell by 1. Depending on the length of the original time step, a periodic system will be encoded as a string similar to

010101010101....

- The decoded output string is then systematically parsed for patterns of a length L , which are collected in a *parse tree*. Figure 2.4a shows the $L = 5$ parse tree for the periodic string.
- The parse tree is further decomposed into reoccurring sub-trees of length D called *morphs* (figure 2.4b). Using the morphs, an ϵ - *machine* is constructed by decoding the full parse tree as an equivalent tree diagram for the morphs (figure 2.4c). The ϵ -machine is called ‘minimal’ if it uses the smallest possible number of morphs.
- Each edge e in the ϵ -machine is assigned a probability p_e showing how likely the system is to proceed along this edge from one morph to another.
- The primary complexity measure S_{SC} from this algorithm is then estimated in analogy to the Shannon entropy S_{Sh} in (2.5):

$$S_{SC} = - \sum_{e=1}^N p_e \log_2 p_e. \quad (2.7)$$

For our perfectly periodic example one thus obtains $S_{SC} = - 2 \times 1/2 \times \log_2(1/2) = 1$, as one might expect from a period system.

The statistical complexity S_{SC} differs from the other complexity measures (section 2.1 and 2.2) in that it assumes maximum values for systems with a large number of equally probable, interconnected morphs. For (truly) chaotic systems, the largest number of such sub-configurations appear at the end of the period-doubling cascade, when the system displays a large number of super-imposed frequencies with some irregularities. This means that for these systems (2.7) really assumes maximum values in the vicinity of “the edge of chaos”, albeit in a parameter space that is not usually a spatio-temporal one.

For example, in their study of a version of the logistic map, Crutchfield (1994) found relatively high statistical complexity values around the critical growth rate r_c , i. e. at the onset of chaos, and lower ones at growth rates further into the chaotic regime. However, this seems to be a particular feature of chaotic maps and their period-doubling cascade to pseudo-randomness. The same paper contains a brief qualitative discussion of the structures found in two simple CAs. It is apparent that an application of the statistical complexity algorithm to these cases will be difficult, and that their maximum values of S_{SC} will not necessarily coincide with the transition between order and disorder. However, the general notion of measuring the amount of structure rather than phase-space distribution as an indicator of complexity corresponds well with the conceptual notions developed in sections 2.1 and 2.2. For this reason Ladyman et al. (2011) endorse S_{SC} as “illustrative of a good measure of complexity” (p. 31).

While statistical complexity has the merit of being computable, and in the limited numbers of examples it has been applied to so far, aligns with intuitive notions of complexity and information distribution, the procedure is practically very unwieldy. The parsing algorithm is computationally expensive and only guarantees pattern recognition up to a chosen length L . Furthermore, any parameterization of the construction of the morphs introduces further estimates (e.g. Crutchfield and Young, 1989; Crutchfield, 1994). In order to allow comparability of different systems, one is further required to find the minimum number of morphs and thus the minimal ϵ -machine representation. These difficulties might be the reason that S_{SC} , despite its emphasize on features traditionally highlighted by complexity research, has not become widely used beyond the confines of Crutchfield’s own research group.

Simplified statistical complexity measures A number of simplified statistical complexity measures were developed in the wake of Crutchfield and Young (1989). These measures also aim to assign vanishing values to both completely ordered and random distributions. Lopez-Ruiz et al. (1995) developed a statistical complexity measure that fulfils these requirements by combining the traditional Shannon entropy in (2.5) with a disequilibrium term D_e . For a set of N outcomes with probabilities p_i they then define statistical complexity $S_{SC,2}$ as

$$S_{SC,2} = S_{Sh} \cdot D_e, \quad (2.8)$$

where

$$D_e = \sum_i \left(p_i - \frac{1}{N} \right)^2. \quad (2.9)$$

D represents the deviation from the uniform distribution. Due to the fact that S_{Sh} will vanish for ordered distributions and the disequilibrium term D_e will tend to zero for random distributions, the measure by definition vanishes for both of these cases. Although a much cruder estimate than S_{SC} , $S_{SC,2}$ clearly has the advantage of simplicity. Several variations of (2.8) have been devised, usually based on a redefinition of the equilibrium value D_e (Martin et al., 2003; Shiner et al., 1999; Feldman

	Random system	Complex system	Chaotic system
<u>Dynamical features</u>			
Size	Large (multi agent system)	Large (multi agent system)	Small (single component system)
Rules	Stochastic	Stochastic or deterministic	Deterministic
Interactions	Undirected (random) interactions	Directed interactions	No interactions
<u>Phenomenological features</u>			
Distribution trends	Random	Random, pseudo-random and patterned	Pseudo-random
Emergent features	-	Structure emergence Complex-to-simple Simple-to-complex	Attractor

Figure 2.5: Comparison of random, complex and chaotic systems.

and Crutchfield, 1998).

2.4 Disentangling complexity from randomness and chaos

2.4.1 A conceptual definition of complexity

Figure 2.5 shows a summary of the features for random, complex and chaotic systems that were highlighted by the discussion above. We have divided the properties into those pertaining to the underlying dynamics and those pertaining to representations of the phase-space output. It is apparent from the table that the only dynamical defining feature, i.e. one that is not shared with either of the two ancestor systems, is the existence of directed interactions between the constituents of the modelled system. ‘Directed’ here means that a constituents only interacts with a certain set of specific other constituents (e.g. its nearest neighbours). A purely dynamical definition would thus be enough to clearly demarcate complex systems:

Definition 1. *Complexity is any manifestation of the directed interactions of a large number of constituents in a multi-agent system.*

However, defining a complex system just as an interconnected multi-agent system would allow systems with homogeneous or periodic phase-space distribution (like the class I and II cellular automata) to be labelled as complex as well. In order to weed out the ‘uninteresting’ cases, we need an additional phenomenological criterion. As we have seen in section 2.2.3, emergence is often named a defining feature of complexity. From figure 2.5 it is again immediately apparent that the distinguishing feature thereby is structure formation. We can thus construct a definition of complexity that neither overlaps with random nor with chaotic systems as follows:

Definition 2. *Complexity is the co-existence of either several patterns, which are clearly distinguishable*

from pseudo-randomness, or at least one pattern and a region of pseudo-randomness, in a suitable representation of a multi-agent system with directed interactions.

Framing this in terms of emergence (under the provision that the emergence categories as defined above are known):

Definition 3. *Complexity is the complex-to-simple or simple-to-complex emergence of one or more patterns, which are clearly distinguishable from pseudo-randomness, in a suitable representation of a multi-agent system with directed interactions.*

It is immediately apparent from the definitions that the phenomenological ‘sieve’ we have included tries to capture the rather elusive and anthropocentric quality that has in various ways been attributed to complexity in the previous sections of this chapter. While there remains an element of subjectivity in this second part of the definition, the dynamical criterion carried over from Definition 1 is a solid and clearly epistemologically useful way of defining the subject of complexity research.

To see how our Definition 2 fares in comparison to other authors’ formal ones, we will consider the following example. Ladyman et al. (2011) set out on a similar quest to ours and defined complexity in the following way (p. 25):

“A complex system is an ensemble of many elements which are interacting in a disordered way, resulting in robust organisation and memory.”

Interaction in a disordered way is clearly a feature of random systems rather than of complex ones (section 2.1 and section 2.2). Further perusal of their article shows that this is in fact just a lapse in language and what they mean is a disordered phase-space distribution. ‘Memory’ and ‘organisation’ have to be similarly gleaned from their article and turn out to mean the development of structures. It remains unclear whether ‘organisation’ refers to phase space appearance or dynamical properties; ‘memory’ is clearly seen as a property of the phase-space portrait. The definition is neither precise enough in distinguishing between dynamics and output nor does it have enough clarity in the use of terminology to meaningfully set aside complexity from chaos or randomness.

Gershenson (2008) is a collection of short interviews with influential complexity researchers. Asked on these questionnaires how they would define complexity, the large majority declines to offer a formal definition. Stephen Wolfram even claims that the existence of a formal definition “tends not to be particularly critical” (p. 136). However, we see several benefits in taking the time to design Definition 2. Firstly, the process of clarifying the boundaries of a concept and disentangling it from its hereditary fields helps to eradicate unsuitable connotations and imprecise language. In our case, the most prominent victim of this pruning process is clearly the ‘edge of chaos’-metaphor. Secondly, the definition itself acts as a guide towards the essential properties of the concept and those that warrant further

Origin	Measure	Maximum for:	Minimum for:	Computability	Lloyd (2001) classification
Statistical mechanics	Boltzmann entropies	Random	Ordered	Computable	Difficulty of description
	Mutual information	-	Random and ordered	Computable	Degree of organization
	Thermodynamic depth	-	Random and ordered	Computable	Difficulty of creation
Computational science	Shannon entropy	Random	Ordered	Computable	Difficulty of description
	Algorithmic information content	Random	Ordered	Incomputable	Difficulty of description
	Logical depth	-	Random and ordered	Incomputable	Difficulty of description
	Effective complexity	-	Random and ordered	Incomputable	Degree of organization
Chaos theory	Statistical complexity		Random and ordered	Computable	Degree of organization

Figure 2.6: Comparison of the complexity measures discussed in section 2.1 to 2.3, including the classification by Lloyd (2001)

investigation. These, in our opinion, are the definition and detection of phase-space patterns as well as the influence of interconnection and interaction in multi-agent systems.

2.4.2 From conceptual definition to complexity measure

Figure 2.6 contains a summary of the different complexity measures discussed in sections 2.1 - 2.3. The conceptual definition of complexity we derived in section 2.4.1 stipulates the existence of patterns in a phase space representation and directed interaction. A suitable quantization of complexity should thus focus on either of these aspects. The existing phase space measures are all based on an analysis of the phase-space output. In agreement with Feldman and Crutchfield (1998) we assume that a practical complexity measure must be computable. This leaves us with a choice between the traditional entropy measures and the statistical complexities. Due to the fact that they are based on the quantification of structures in the phase space output, we consider the statistical complexity a better tool to measure complexity as we have defined it.

In the following two chapters we will put our choice to the test and examine three simple systems - one complex, one random and one chaotic - with both the Shannon entropy S_{Sh} and the simplified statistical complexity $S_{SC,2}$. We will then use the results of this comparison and our conceptual definition of complexity to sketch out some directions for the development of prospective complexity measures.

Chapter 3

Three models

In this chapter we will introduce three simple models: the random Kac ring model (section 3.1); the complex Bak-Sneppen evolution model (section 3.2) and the chaotic logistic equation (section 3.3). All three models have been extensively studied on their own but so far there exist no comparative study.

All the computer models were programmed in MathCad (appendix A1) and iterated for 200 time steps. Appendix 2 contains a number of sample animations illustrating how these systems evolved. The number of runs, the integration time and the number of constituents (which will be motivated below for each particular model) are at the lower end of what would be acceptable for an exploratory computer study. However, since we can often verify our results directly against previous work and since the modelling work mainly functions to illustrate the conceptual results of the thesis, we consider them fully adequate for the present study.

3.1 A random system: The Kac ring

The Kac model was developed by Kac (1956) as a means of demonstrating that coarse graining and *Stosszahlansatz*-like dynamical assumptions can lead to irreversible evolution of a macroscopic entropy. It describes an idealized system of particles whose phase-space values are determined by randomized collisions only. Dorfman (1999) and Gottwald and Oliver (2009) contain a detailed discussion of the dynamics of the system in the context of thermodynamics and statistical physics. Our abbreviated account below is based on their work.

3.1.1 Dynamics of the Kac ring

The KAC model consists of a one-dimensional, periodic lattice of N sites arranged around a ring. Each site is occupied by a particle, which has either the property ‘black’ or the property ‘white’. Neighbouring sites are joined by edges. Each of the N edges is assigned one of two markers, which will be denoted as S and \bar{S} .

The dynamics of the system are discrete and consist of particles moving clockwise to the neighbouring site. Thereby, the edge markers control the evolution of the colour property: a particle changes colour if it traverses an edge marked \bar{S} but remains unchanged if the edge marker is S . If the same number of steps is retraced in the counter-clockwise direction, the initial state is reached again. The dynamics of the ring are therefore reversible. Furthermore, the time evolution of the system is periodic. Depending on the number n of colour changing edges \bar{S} , states reoccur with a period of N or $2N$ time-steps for even or odd n , respectively.

Mimicking the dual mode of description employed in statistical physics and thermodynamics, microscopic and macroscopic variables will be used to describe the KAC model. On the macroscopic level, the number B of black particles and the number W of white particles will be used as fundamental variables. On a mesoscale level, convenient properties to consider are the number of black or white particles before an \bar{S} edge: b and w , respectively. From the previous paragraph, the importance of these quantities for the evolution of the model in the next step is immediately apparent: b is equivalent to the number of particles changing from black to white and w is equivalent to the number of particles changing from white to black. Since the computation of b and w requires knowledge of the exact location of the particles on the ring, while that of B and W only relies on the overall ratio of particles regardless of their location, the choice of these two pairs of properties is in keeping with analogy to a microscopic/macroscopic description in statistical mechanics.

Using these definitions, the dynamics of the macroscopic quantities are given by the following equations:

$$B(t+1) = B(t) + w(t) - b(t) \tag{3.1}$$

and

$$W(t+1) = W(t) - w(t) + b(t), \tag{3.2}$$

where t is the discrete time step and constitutes a single, clockwise shift of the ring. (3.1) and (3.2) can be combined into a single difference equation:

$$\Delta(t+1) = B(t+1) - W(t+1) = \Delta(t) + 2w(t) - 2b(t). \tag{3.3}$$

Since equation (3.3) still depends crucially on the microscopic properties b and w , no closed macroscopic description has been achieved. In order to eradicate these variables one needs to make further assumptions about their relation to the large scale properties of the system. In analogy to Boltzmann's *Soszahlansatz* (section 2.1.1), one assumes that any particle is equally likely to experience a 'collision' and change their colour property. Noting that the probability of an edge around the ring carrying the

a)

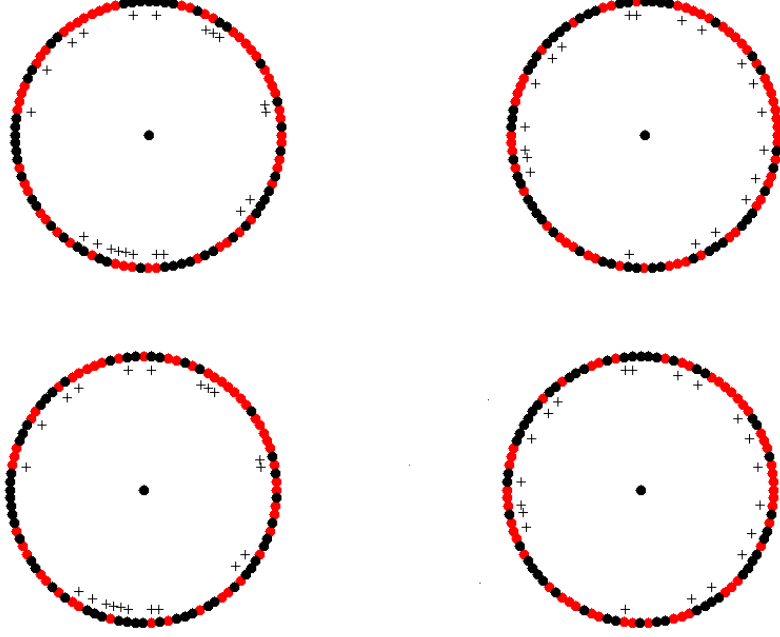


Figure 3.1: Examples of the Kac ring simulations. The left and right row show two different runs at time-steps 150 and 151 each. An x indicates the location of an \bar{S} edge.

marker \bar{S} is given by n/N , this means one assumes that this ratio must also be the probability for a particle to change colour, i.e. one requires:

$$\mu := \frac{n}{N} = \frac{b}{B} = \frac{w}{W}. \quad (3.4)$$

Since μ can be derived from the macroscopic quantities n and N , we can use (3.4) to remove the microscopic quantities from (3.3):

$$\Delta(t+1) = \Delta(t) + 2\mu(W(t) - B(t)) = \Delta(t)(1 - 2\mu). \quad (3.5)$$

(3.5) defines a geometric sequence, which can be rewritten as a function of the initial state $\Delta P(0)$.

$$\Delta(t) = (1 - 2\mu)^t \Delta(0). \quad (3.6)$$

It is notable that (3.6) is not time reversible, since we necessarily have $\mu \leq 1$ and therefore $\Delta(t) \rightarrow 0$ for $t \rightarrow \infty$. The only exception to this asymptotic behaviour is an initial set-up where the number of S edges equals the number of \bar{S} edges and thus $n = (1/2)N$, in which case (3.6) is trivially zero. The Kac model thus captures the time development contrast between macroscopic and microscopic variables that is characteristic of statistical mechanical systems.

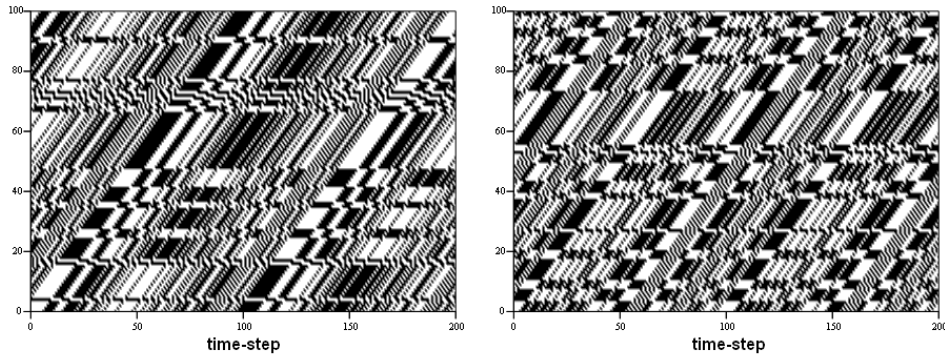


Figure 3.2: Phase-space diagrams of the Kac ring for the two runs in figure 3.1.

3.1.2 Simulations of the Kac ring

We have simulated a Kac ring of 101 sites and fixed $\mu = 0.2$. The initial state was a uniformly random distribution of black and white sites around the ring. From this initial state an ensemble of 20 runs was computed whereas the distribution of edges \bar{S} was determined anew for each run. Figure 3.1 shows an example of how different runs develop at two consecutive time-steps. Phase-space diagrams of the same runs are shown in figure 3.2.

Figure 3.3 shows the development of $\Delta(t)$ for the ensemble of 20 runs. On average $\Delta(t)$ shows only small fluctuations around zero, as expected from (3.6). A control run was conducted (appendix A2) with an initial state of only white sites, which verified that the model would tend towards a parity distribution even if started of in an extreme deviation from the equilibrium state. Gottwald and Oliver (2009) proved that the relative Δ -variance of an ensemble of Kac rings scales with $1/\sqrt{N}$. It is therefore not surprising that our collection of small rings shows a relatively large variance.

3.2 A complex model: The Bak-Sneppen model

The Bak-Sneppen model was developed by Bak and Sneppen (1993). It has since become a much-cited example for a system with self-organized criticality (e.g. Bak and Paczuski, 1995; Frigg, 2003; Kauffmann, 1995) and thus holds at least some epistemological credentials towards a claim of complexity (section 2.2.3). The natural process the model represents is the co-evolution of a number of species.

3.2.1 Dynamics of the Bak-Sneppen model

The original dynamics of the Bak-Sneppen model were stochastic and had a continuous range of output values between 0 and 1 (Bak and Sneppen, 1993). In order to compare with our other two models, we have included two later versions as well, a range-discrete model (Meester and Znamenski, 2002) and a deterministic model (e.g. de Los Rios et al., 1997; Mendes, 2006).

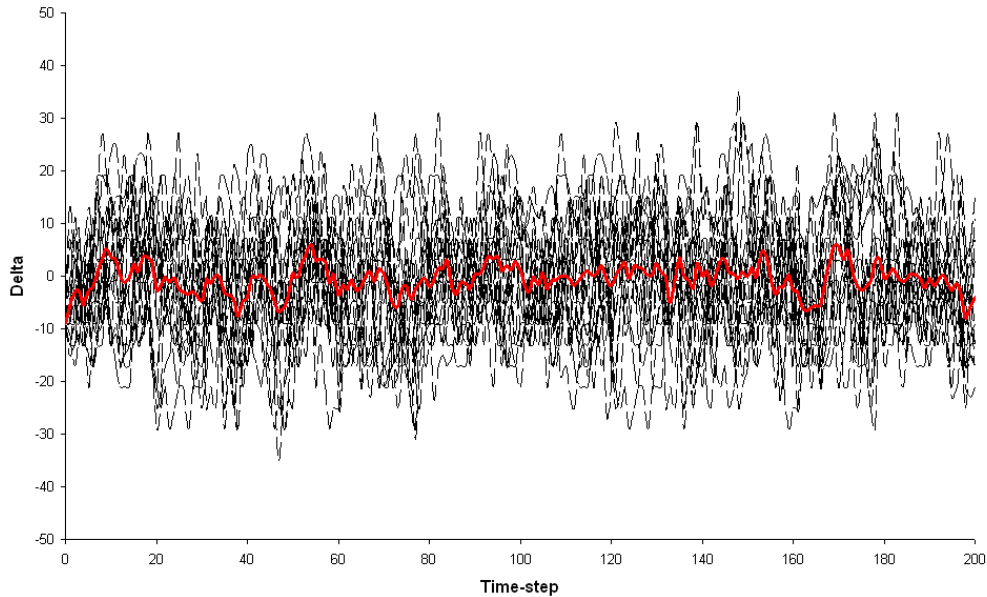


Figure 3.3: $\Delta(t)$ for the ensemble of Kac rings. The ensemble average is plotted in red.

Classic Bak-Sneppen model In its one-dimensional form, which we will consider for better comparison to the other two models discussed, the Bak-Sneppen model consists of a ring of N sites connected to each other by edges. The general set-up is therefore very similar to the Kac model. However, instead of being governed by the properties of the edges, the time evolution is controlled by a set of rules prescribing the interaction of neighbouring sites. Each site represents a species with fitness value x_i ($x_i \in [0,1]$). The development of connected sites - here the triple x_{i-1} , x_i and x_{i+1} - is interlinked, in a crude attempt to model the evolution of co-dependent species.

During a time-step, the site with the lowest fitness value x_{min} is determined. Its fitness value and those of its two neighbours, x_{min-1} and x_{min+1} , are then replaced by equi-distributed random numbers on the unit interval.

Deterministic Bak-Sneppen model The general set-up of the deterministic Bak-Sneppen model is identical to the stochastic version. However, instead of updating x_{min} and its neighbours by random numbers, a deterministic function is employed. In previous studies, deterministic functions that are sufficiently variable led to output that was virtually indistinguishable from the classic model (e.g., for review, Gillet, 2007). Following a study by Mendes (2006), we employ the deterministic updating rule:

$$x_{i+1} \rightarrow kx_i \text{ mod } 1, \quad (3.7)$$

where k is a constant that can be freely chosen. All results shown here were obtained for $k = 2$.

Discrete Bak-Sneppen model A slightly modified version of the model uses fitness values of 0 and 1 only. At each time step one site x_{rand} is chosen randomly. If $x_{rand} = 0$ then the triple around x_{rand}

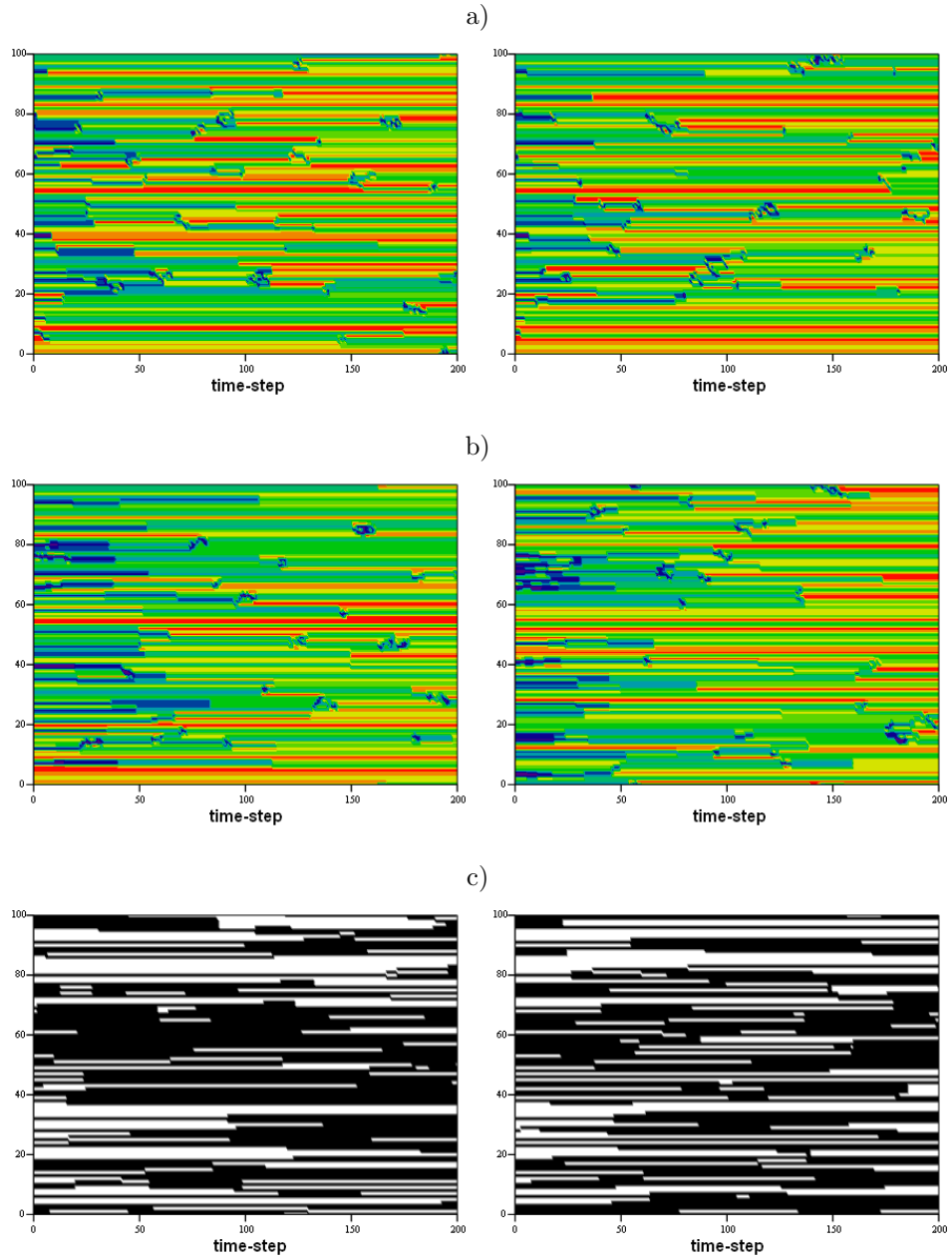


Figure 3.4: Phase-space portraits of the three Bak-Sneppen simulations. a) Classic model. b) Deterministic model. The colour scheme goes from violet, indicating zero, to red, indicating one. c) Discrete model (black indicates a fitness value of 1, white indicates a fitness value of 0).

is updated by a renewed fitness calculation. This computation assigns a new fitness value of 1 with a probability p and a new fitness value of 0 with a probability of $1 - p$ to each site. All models shown here employed $p = 0.2$.

3.2.2 Simulations of the Bak-Sneppen model

We have simulated ensembles of 20 runs each for the three versions of the Bak-Sneppen model. The number of constituents was $N=101$. The two stochastic ensembles were started from the same uniformly random initial distributions. In order to obtain an ensemble for the deterministic case, a different uniformly random initial distribution was generated for every run.

In their pioneering study, Bak and Sneppen (1993) described the dynamics of the classic Bak-Sneppen model as follows (p. 4085):

“We [...] observe long periods of passivity interrupted by sudden bursts of activity: the model exhibits punctuated equilibria.”

The equilibrium values towards which the system organizes itself has been determined to be approximately $r_c = 2/3$ (e.g. Gillet, 2007; Grassberger, 1995). The episodes of punctuated equilibrium are identical with the ‘avalanches’ discussed in section 2.2.3.

Figure 3.4 shows the phase-space diagrams for the three Bak-Sneppen models. A comparison with figure 2.2 immediately shows that, for all three models, we observe the expected punctuated equilibrium dynamics. Similarly, we also observe an overall tendency of the continuous versions (figures 3.4a and 3.4b) to approach r_c . This trend is best illustrated in the sample animations include in appendix A2. Despite the fact that the technical limitations and the focus of our study render it infeasible to verify the power law distribution of avalanches associated with self-organized criticality, we are confident that our models adequately represent the dynamics of the Bak-Sneppen model.

Qualitative complexity of the Bak-Sneppen model We will need to satisfy ourselves that the Bak-Sneppen models and the phase-space representation in figure 3.4 are indeed complex according to Definition 2 (chapter 2). This will also serve as an example of an application of the definition. The models clearly fulfil the requirements on their dynamics of being ‘multi-agent system[s] with direct interactions’. In contrast to the Kac ring discussed above, interaction in the model are not parameterized by a random algorithm but by nearest neighbour dynamics, which specify that a site only interacts with its two closest neighbours.

The second, phenomenological part of the definition would not be needed to demarcate the Bak-Sneppen model from those models discussed in sections 3.1 and 3.2; the dynamical part of Definition 2 is sufficient for this. However, proving that the model fulfils this part as well, will support claims that it can be seen as a representative example of a complex model. Definition 2 requires the identifications of patterns in a representation of the system. The complex quantity of figure 3.4 arises from the co-existence of the interrupted stripe patterns with the irregular ‘avalanche’-pattern. In this figure 3.4 clearly differs from figure 3.2, which shows nested arrangements of a block pattern, but no additional structural features beyond this. As suspected in section 2.4.2, the phenomenological identification of a system as complex turns out to be less clear-cut than the dynamical one.

The Bak-Sneppen model has been discussed as an example in the context of complex-to-simple emergence (e.g Kauffman, 1995). As described in section 2.2.3, the appearance of the avalanches is thereby interpreted as a self-regulation mechanism with which the system steers itself towards the critical value

r_c . Neither the appearance of this mechanism nor the asymptotic behaviour are trivially predictable from the stochastic/pseudo-random dynamics. This categorization crucially hinges on the fact that the rules of the Bak-Sneppen model are judged to be more complex than its output. A reversal of this attribution of simplicity and complexity could conceivably be achieved by focusing on the conciseness with which these rules can be stated rather than on the generation of random/pseudo-random numbers in the algorithm, and by contrasting it with the existence of two different, complicated patterns in the phase-space diagram.

3.3 A chaotic model: The logistic equation

In what is generally perceived to be the first study of chaos in biological systems, May (1974, 1976a,b) conducted an in-depth examination of the discrete forms of the logistic equation,

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K}\right). \quad (3.8)$$

The equation describes the development of the population numbers P with time t for a single species with growth rate r . Clearly (3.8) has a global, stable equilibrium at $P = K$. This means that once a species has reached a population number of K , it will stay at this size. Following May (1974)'s discovery of its chaotic regime, the logistic equation has become a prototype of chaotic maps and was the basis for many influential discoveries in the field (e.g. Feigenbaum, 1977; Grassberger and Procaccia, 1983).

3.3.1 The dynamics of the logistic equation

The discrete form of (3.8) is given by:

$$P_{t+1} = P_t \left(1 + r \left(1 - \frac{P_t}{K}\right)\right). \quad (3.9)$$

Since (3.9) allows negative values for the population numbers, May (1974) focused their research on a slightly altered version which treats the right hand side of (3.9) as the first two terms of a Taylor expansion and contracts into

$$P_{t+1} = P_t \exp \left(r \left(1 - \frac{P_t}{K}\right)\right). \quad (3.10)$$

3.3.2 Simulations of the logistic equation

The logistic equation has very different dynamics from the two models discussed in sections 3.1 and 3.2: it is a single-component system whose behaviour is governed by a recurrence relation with independent parameter r . As announced in chapter 1, in order to achieve a meaningful comparison to the other two models, we need to construct a system similar to a multi-agent system but composed of non-interacting constituents. We do this by considering a set of $N=101$ logistic equations (3.10) on the parameter space $r \in [0,5]$. In all the following representations, the equations are ordered so that the i -th element of this set

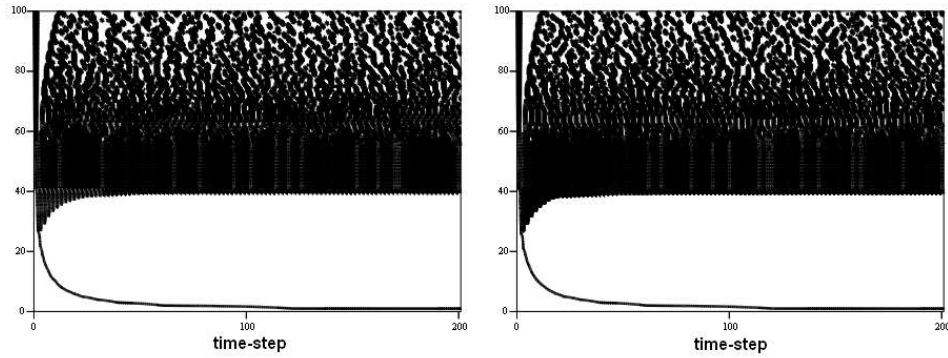


Figure 3.5: Phase-space diagrams for the logistic equation

has a growth rate of $r_i = 0.2 \cdot i$. This way of representation differs from the usual mode of investigation of chaotic systems (e.g. Hilborn, 2002), which focuses strongly on the recurrence properties of the maps. However, it is easily possible to construe a natural situation that corresponds to our model of several logistic equations: if the Bak-Sneppen model describes the co-evolution of N species, then a collection of logistic equations describes the parallel evolution of N unrelated species with different growth rates r .

We simulated an ensemble of 20 such sets of logistic equations, with initial values of $P_0 = 0.1$ and equilibrium value of $K = 0.5$ in normalized units. Integrating (3.8) with time leads to:

$$P(t) = \frac{K}{1 + \frac{1-P_0/K}{P_0/K} \exp(-rt)}, \quad (3.11)$$

where P_0 is the initial population value. It is immediate apparent from (3.11) that all populations will eventually approach the equilibrium population K . The rapidity with which this process happens depends on the growth rate r .

In their analysis of (3.10), May (1974) noted that the discrete population equation shows widely different behaviour in dependence on the parameter r . They divided the solutions they obtained into three different dynamical regimes - an equilibrium regime, a periodic regime and a chaotic regime. For small growth rates ($r < 2$) the behaviour of the discrete solutions mimics that of the analytic ones. Depending on the growth rate, all solutions eventually approach the equilibrium population K . Accordingly, the parameter space for these small values of r has been named the ‘equilibrium regime’.

If r falls between 2.0 and the critical value $r_c = 2.692$, the iterated solutions oscillate around the equilibrium value K . Scanning through the parameter space towards r_c , one observes period doubling, e.g. more oscillations with separate peaks are interposed over the original one. The first doubling takes place at $r = 2.526$, the second one at $r = 2.656$ and so on. This is called the *period-doubling cascade* to chaos (section 2.3.3)

For growth rates beyond the critical value, the solutions become erratic, displaying either quasi-periodic oscillations or seemingly random outbursts of activity (May, 1976a). Due to the fact the solutions now defy forward extrapolation, this regime is called the ‘chaotic regime’. The qualitative development is the same for both (3.9) and (3.10) but the critical values for the regime transitions are slightly different.

Figure 3.5 shows the phase-space diagrams for our set of logistic populations. The three regimes are easily discernible. The transition from uniformity to chaos via periodicity is also visible in the animation in appendix A2, although our resolution in r space is not fine enough to resolve much of the period doubling cascade. However, our models are clearly in good qualitative agreement with previous results.

Phenomenological similarity to complexity Our set of logistic equations is not dynamically complex, since it lacks the required directed interactions. However, the phase-space diagrams in figure 3.5 show a phenomenological similarity to complexity. They combine structure, in the form of homogeneity and periodicity with pseudo-randomness. Without knowledge of the underlying dynamics, one would thus not be able to use Definition 2 (section 2.4.2) in order to label the Bak-Sneppen model as complex and the logistic system as chaotic.

It is possible to revise Definition 1 to phenomenologically exclude chaotic systems. For example, relying on the fact that period-doubling is probably a universal route to chaos (e.g. Feigenbaum, 1977) one could insert an explicit exclusion of periodicity and so require:

Definition 4. *Complexity is the co-existence of either several patterns, which are clearly distinguishable from pseudo-randomness, periodicity and homogeneity, or at least one pattern and a region of pseudo-randomness, in a suitable representation of a multi-agent system with directed interactions.*

This would effectively demarcate between figures 3.4 and 3.5, but will also filter out a large number of those systems that fulfil the dynamical complexity criterion.

Another way to achieve a phenomenological distinction is the establishment of stringent requirements for suitable representations and thus the exclusion of 3.5. This clearly leads back to our discussion of the distinction between chaos and pseudo-randomness (chapter 2) and the central definition of chaos as the property of maps rather than multi-component systems (chapter 1). Since the elements in our set of logistic equations are completely independent and thus could be plotted in arbitrary order, we could reasonably argue that 3.5 is only formally equivalent to 3.4.

Chapter 4

Comparing complexity: Shannon entropy and statistical complexity

In the present chapter we will compute estimates of the Shannon entropy and the statistical complexity for the five models described in chapter 3. As discussed, the statistical complexity is the measure we consider best matched to our qualitative definition of complexity (section 2.4.2). The Shannon entropy (section 2.2.4) is a representative of the traditional entropy and it is normalized, which makes it a better means of comparison than its close cousin, Boltzmann's combinatorial entropy (section 2.1.1). There are several technical difficulties that arise when using such measures, most prominently the choice of a suitable phase-space partition. A comprehensive treatment of these is beyond the scope of this thesis and we will content ourself with the knowledge that the technical means employed are for the most part conventional ones.

4.1 Complexity estimates for the Kac ring

Shannon entropy of the Kac ring Gottwald and Oliver (2009) computed one second concept of Boltzmannian entropy, $S_{B,2}$, for the Kac ring model. They assume the ring's phase space to consist of particles' colour states x_i , which can be either white or black. The positions of the particles along the ring can be neglected, since the set-up is perfectly symmetric. A natural partition for this space is two assume two cells only, one containing all black particles and one containing all white particles. $S_{B,2}$ can then be computed from (2.4):

$$S_{B,2}^{Kac} = \frac{N!}{B(t)!W(t)!}. \quad (4.1)$$

The Shannon entropy can be defined in analogy to this by computing the probabilities for a particle to be found in the black or white box, respectively, as:

$$p_B(t) = \frac{B(t)}{N} \quad (4.2)$$

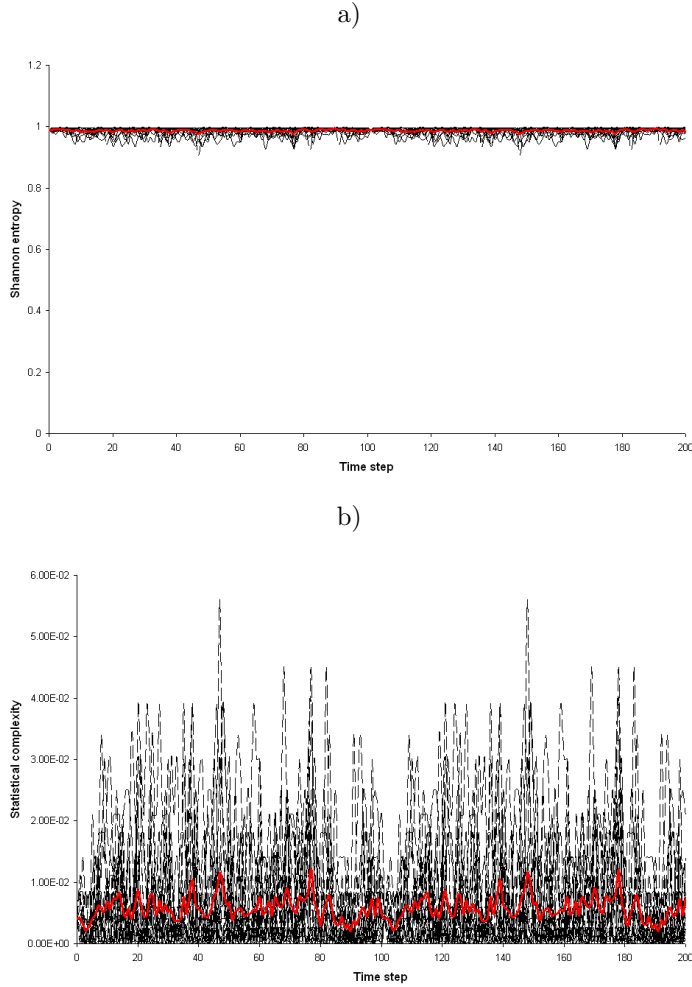


Figure 4.1: a) Shannon entropy S_{Sh} and b) statistical complexity $S_{SC,2}$ for the ensemble of Kac rings. The ensemble average is plotted in red.

$$p_W(t) = \frac{W(t)}{N}. \quad (4.3)$$

Using (2.5) one then obtains:

$$S_{Sh}^{Kac} = p_B \log_2(p_B) + p_W \log_2 p_W. \quad (4.4)$$

Figure 4.1a show the Shannon entropy for the ensemble of Kac rings. As expected for a thermodynamic system, all models quickly approaches the maximum state of a random distribution and then stay at maximum entropy for the remainder of the run. This is consistent with the results obtained by Gottwald and Oliver (2009) and the $\Delta(t)$ development in figure 3.3.

Statistical complexity of the Kac ring We can compute the (simplified) statistical complexity for the Kac ring by inserting (4.4), (4.2) and (4.3) into (2.8):

$$S_{SC,2}^{Kac} = S_{Sh}^{Kac} \left[\left(p_B - \frac{1}{2} \right)^2 + \left(p_W - \frac{1}{2} \right)^2 \right]. \quad (4.5)$$

Figure 4.1b shows the statistical complexity for the ensemble of Kac rings. On average, the complexity is close to zero and constant. This is due to the minimization of the disequilibrium factor in (4.5)

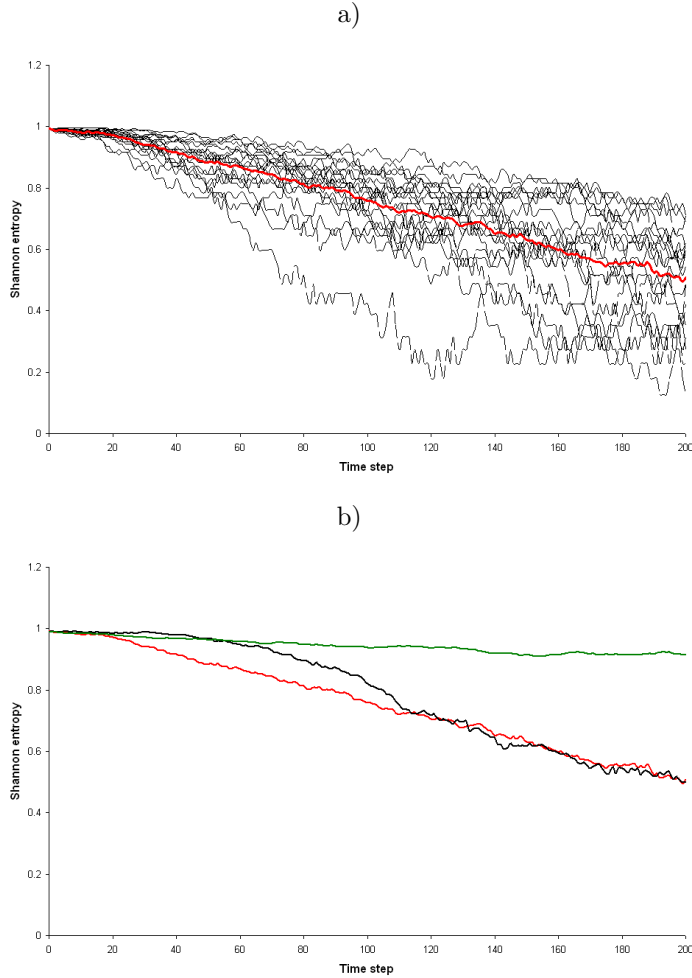


Figure 4.2: a) Shannon entropy S_{Sh} for the ensemble of classic Bak-Sneppen models. The ensemble average is plotted in red. b) Average Shannon entropy S_{Sh} for the three Bak-Sneppen models. Black: Classic. Red: Deterministic. Green: Discrete.

through the tendency of the model to assume an equal distribution between the two phase space cells.

4.2 Complexity estimates for the Bak-Sneppen model

Shannon entropy of the Bak-Sneppen model The phase space of the Bak-Sneppen model is given by the N -dimensional vector of fitness values $X = (x_1, \dots, x_N) \in [0, 1]^N$. We are then again faced with the challenge of choosing a meaningful partition.

- In the range-discrete model the partition is trivially give as a division between fitness values 1 and 0. This yields for the Shannon entropy:

$$S_{Sh}^{BS,dis} = \frac{N_0}{N} \log_2 \left(\frac{N_0}{N} \right) + \frac{N_1}{N} \log_2 \left(\frac{N_1}{N} \right), \quad (4.6)$$

where N_0 is the number of species with fitness value 0 and N_1 is the number of species with fitness value 1.

- In the classic and the deterministic Bak-Sneppen model, the fitness values can take any value

between 0 and 1. Since the values are chosen from an uniform distribution, a possible tactic is to divide the space at mid-point, creating a partition $([0, 0.5], [0.5, 1])$ that could be roughly interpreted as classifying species as strong ($x > 0.5$) and weak ($x < 0.5$). The corresponding Shannon entropy is given by

$$S_{Sh}^{BS} = \frac{N_{<0.5}}{N} \log_2 \left(\frac{N_{<0.5}}{N} \right) + \frac{N_{>0.5}}{N} \log_2 \left(\frac{N_{>0.5}}{N} \right), \quad (4.7)$$

where $N_{>0.5}$ is the number of strong species and $N_{<0.5}$ is the number of weak species.

- To demonstrate that the choice of coarse graining is somewhat arbitrary, we could consider the following alternative partition: knowing that the classic Bak-Sneppen models tend to approach a critical fitness value r_c (section 3.2), another natural partition would be into species below and above r_c , $([0, r_c], [r_c, 1])$. The Shannon entropy can then be calculated in analogy to (4.7).

Figure 4.2a shows the bi-partitioned Shannon entropy time-development for the ensemble of classical Bak-Sneppen models. The models show an increasingly large variance as the run progresses but the ensemble average is still a good estimate for the trend in the models. This is true for the discrete and deterministic models as well (results not shown). For all three models, the overall trend is a continuous decrease in Shannon-entropy as the run progresses (figure 4.2b). The models are therefore becoming more ordered. This can be attributed to the approach of the critical value r_c , which we observed in chapter 3. Accordingly, the phase-space cell of weaker species is depleted and the system moves towards confinement in the upper cell alone. The directed development of the Bak-Sneppen models, as opposed to the random one of the Kac ring, is thus well captured by the entropy profile.

Statistical complexity of the Bak-Sneppen model The statistical complexity can be computed from (4.6) and (4.7) by multiplication with disequilibrium term (2.9), yielding:

$$S_{SC,2}^{BS,dis} = S_{Sh}^{BS,dis} \cdot \left[\left(\frac{N_0}{N} - \frac{1}{2} \right)^2 + \left(\frac{N_1}{N} - \frac{1}{2} \right)^2 \right] \quad (4.8)$$

and

$$S_{SC,2}^{BS} = S_{Sh}^{BS} \cdot \left[\left(\frac{N_{<0.5}}{N} - \frac{1}{2} \right)^2 + \left(\frac{N_{>0.5}}{N} - \frac{1}{2} \right)^2 \right]. \quad (4.9)$$

Figure 4.3a shows the statistical complexity for the ensemble of classical Bak-Sneppen models. It is obvious from the figure that the average is a good representative of the general trends. In contrast to the entropy measurements, the statistical entropy increases throughout the runs and then levels out towards a maximum value. The overall development is clearly dictated by the approach of the critical value r_c . In the single profiles one can also see the signature of the avalanches as very sharp dips in the profile. All three models show the same trends in complexity development (figure 4.3b). Thereby the classic and deterministic model are again very similar, while the discrete model progresses much more slowly.

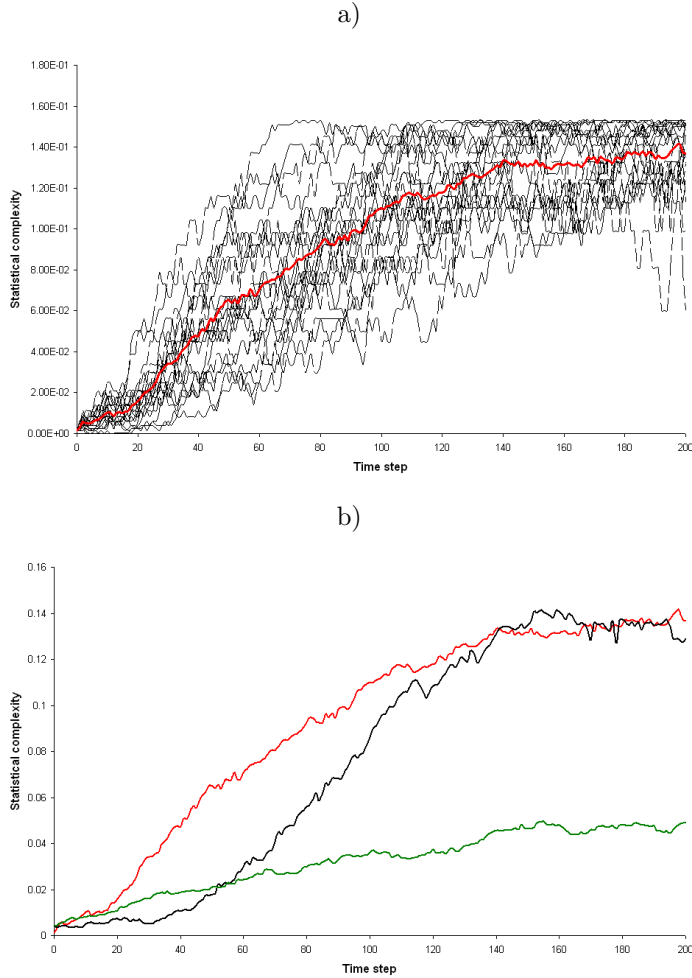


Figure 4.3: a) Statistical complexity $S_{SC,2}$ for the ensemble of classic Bak-Sneppen models. The ensemble average is plotted in red. b) Average statistical complexity $S_{SC,2}$ for the three Bak-Sneppen models. Black: Classic. Red: Deterministic. Green: Discrete.

4.3 Complexity estimates for the logistic equation

Shannon entropy of the logistic equation The phase-space in the set of logistic equations we discussed in section 3.3 is given by the population numbers P_i , $X = (P_1, \dots, P_N) \in [0, 1]^N$. Choosing a partition is again difficult. However, we now have the precedent of Crutchfield and Young (1989), who used a partition at the mid-point ($[0, 0.5), [0.5, 1]$). In our case this involves dividing the phase-space into a cell above and one below the equilibrium value. The resulting Shannon entropy is then given by

$$S_{Sh}^{Log} = \frac{N_{<Eq.}}{N} \log_2 \left(\frac{N_{<Eq.}}{N} \right) + \frac{N_{>Eq.}}{N} \log_2 \left(\frac{N_{>Eq.}}{N} \right), \quad (4.10)$$

where $N_{<Eq.}$ is the number of population values below the equilibrium values and $N_{>Eq.}$ is the number of population values above the equilibrium value. In previous studies complexity measures were calculated as a function of the growth rate r (Crutchfield and Young, 1989; Lopez-Ruiz et al., 1995). In this case $N_{<Eq.}$ and $N_{>Eq.}$ are computed from the time profile, i.e. (3.10) is iterated for some time and then sorted into the phase-space partition. However, in our case we have set up the system of populations in analogy to the Kac ring and the Bak-Sneppen models and are therefore mostly interested in the

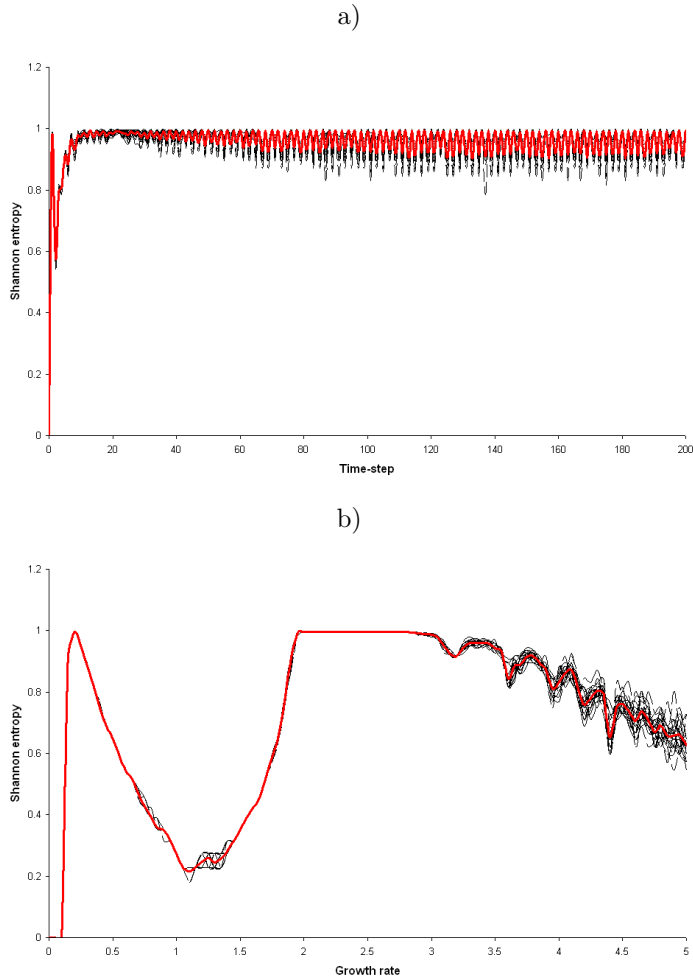


Figure 4.4: Shannon entropy S_{Sh} for the logistic equation a) plotted with time and b) plotted with growth rate. The ensemble average is shown in red.

time-development of the whole set of equations (section 3.3.1). We will therefore also compute (4.10) for the set of equations with different r -values.

Figure 4.4a shows the Shannon entropy for the system of logistic populations. Similarly to the Kac-ring model, the system quickly approaches a maximum value. This is clearly due to the periodic and pseudo-random fluctuations shown by the majority of the constituents in the system (section 3.3.2). Both of these lead to an approximately even spread through the two phase-space cells. In contrast to the other two models, the system shows rapid oscillations around the entropy value it approaches.

The Shannon entropy with growth rate r is plotted in figure 4.4b. The entropy development clearly picks out the different regions of the parameter space. Low values are assigned to the equilibrium regime while the periodic regime leads to a maximum spread through the two phase-space cells, and thus maximized the entropy. The entropy is generally high in the chaotic regime, where one sees pseudo-random fluctuations. The more pronounced relative maxima seem to correspond to the ‘ordered windows’ in the chaotic regime, i.e. regions where the system approaches periodicity again (e.g Hilborn, 2002).

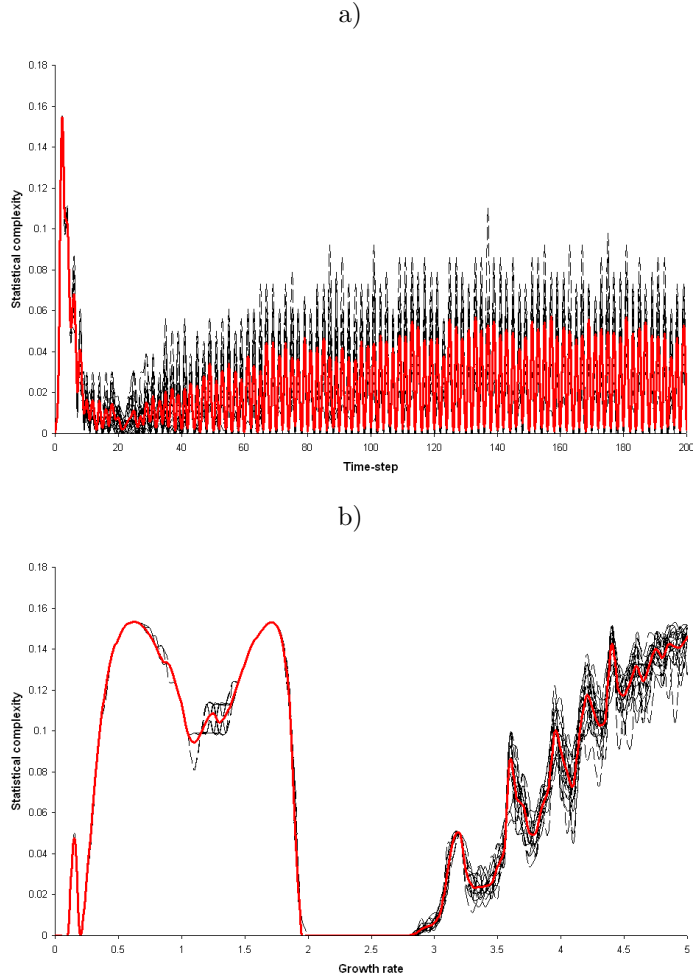


Figure 4.5: a) Statistical complexity $S_{SC,2}$ for the logistic equation a) plotted with time and b) plotted with growth rate. The ensemble average is shown in red.

Statistical complexity The statistical complexity for the logistic equation can be defined from (4.10) and (2.8):

$$S_{SC,2}^{Log} = S_{Sh}^{Log} \cdot \left[\left(\frac{N_{<0.5}}{N} - \frac{1}{2} \right)^2 + \left(\frac{N_{>0.5}}{N} - \frac{1}{2} \right)^2 \right]. \quad (4.11)$$

We again compute (4.11) with time for the whole system of equations or as a function of the growth rate r .

Figure 4.5a shows the statistical complexity of the logistic systems as a function of time. The complexity is much lower than the entropy and shows even larger small scale fluctuations. The disequilibrium value is low for most of the system's constituents since the majority of r values displays pseudo-random oscillations. There are also large fluctuations within the ensemble. In this, as expected, the chaotic system is very similar to the Kac ring, which shows truly random behaviour.

The results of the complexity development with growth rate r , which are displayed in figure 4.5b also show the three distinct parameter regimes. The fact that both the complexity and the Shannon entropy have finite (instead of zero as we would expect) values in the equilibrium regime seems to be

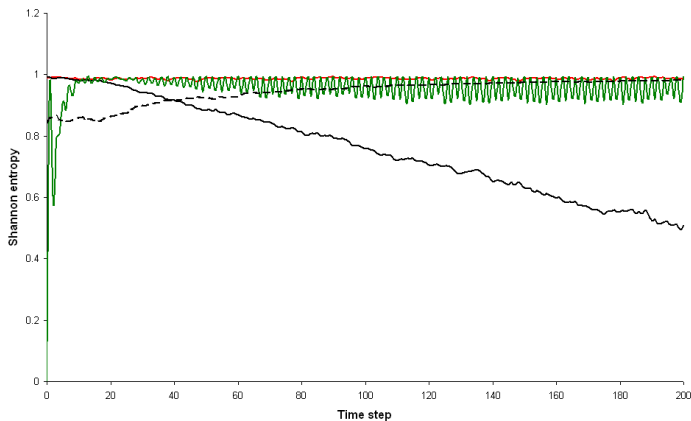


Figure 4.6: Average Shannon entropy S_{Sh} for all models. Black: Classic Bak-Sneppen model. Red: Kac ring. Dashed-Black: Bak-Sneppen model with phase-space division at $r_c = 2/3$.

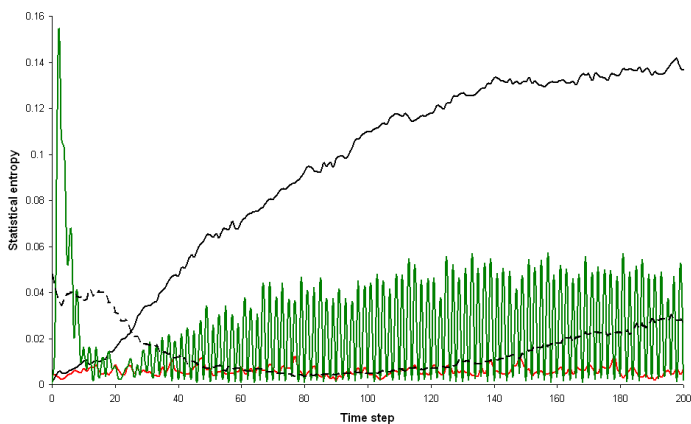


Figure 4.7: Average Shannon entropy S_{Sh} for all models. Black: Classic Bak-Sneppen model. Red: Kac ring. Dashed-Black: Bak-Sneppen model with phase-space division at $r_c = 2/3$.

an artefact of our model's set-up and the phase-space partition we use: in the early stages of the run the models often show minute but irregular fluctuations around the equilibrium value. This can lead to both finite S_{Sh} and a finite D_e . In the periodic regime, the complexity is zero. This is due to the fact that, in our phase-space partition, oscillations around the 0.5 equilibrium value will lead to equal numbers in both cells (e.g. Lopez-Ruiz et al., 1995). As observed by Crutchfield and Young (1989), the complexity rises sharply at the end of the period-doubling cascade. It keeps rising throughout the chaotic regime, a results that could not have been observed and was not anticipated by Crutchfield and Young (1989), who only investigated a small parameter region around the critical value. The intermittent maxima and minima in this region again correspond roughly to the major bands of irregular outbursts and quasi-periodicity, respectively. However, our results do suffer from the fact that we are looking at the system through a very low r -resolution and further studies are needed to conclusively confirm them.

4.4 Comparison of the models' complexity

Shannon entropy Figure 4.6 shows a comparison of the average Shannon entropy profiles for the Kac ring, the classic Bak-Sneppen model and the logistic system. The absolute highest entropy values are reached for the Kac ring, followed by the logistic system, which also distinguishes itself through the rapid oscillations still visible in the average profile, and then the Bak-Sneppen system. The complex system Bak-Sneppen model is clearly set apart from the chaotic and the random one by being the only one with declining entropy values. As pointed out in section 4.2, this is due to the model's tendency towards the critical value $r_c = 2/3$, which in our partition is placed in the upper phase-space cell. This has the disturbing consequence that if we realise the alternative partition considered in section 4.1 and divide the phase-space at r_c , then the entropy profile for the Bak-Sneppen model changes significantly and the model now also approaches a maximum value (figure 4.6). In this scenario, all three models show converging entropies and we would be hard pressed to distinguish the complex model from the two others.

Statistical complexity The average statistical complexity for all three models is shown in figure 4.7. If the complexity is calculated for all models on the bi-partition, then the Bak-Sneppen model is clearly marked as the most complex. The Kac ring is clearly of low complexity. The logistic equation has a low general complexity value but shows wild small-scale oscillations. For this particular phase-space partition, the statistical complexity performs very well and clearly marks out the complex system. Unfortunately, if the r_c -partition is chosen, the Bak-Sneppen system has a much lower statistical complexity and even falls below the logistic equation.

For both S_{Sh} and $S_{SC,2}$ the observations described in the previous two paragraphs are also valid if single runs rather than ensemble averages are compared.

Difficulties of comparison The discussion of this and the previous chapter have highlighted certain difficulties with the use of both entropy and complexity measures to compare different systems. First of all, it is necessary to either construct (in a theoretical study) or find (in nature) systems that are actually comparable. As we have seen in section 3.3.1, this is not an easy task. Secondly, one is faced with the difficulty of defining a phase-space partition that is both natural to each system and preserves comparability. As we have seen in figures 4.6 and 4.7, the choice of partition can greatly alter the qualitative nature of the entropy and complexity results. This aspect has received very little attention from the current literature on complexity measures (e.g Feldman and Crutchfield, 1998; Bennett, 2003). In fact, both complexity measure seem to perform best in characterising different phases of a given systems development rather than yielding themselves to inter-system comparison. This might be the reason that the existing studies have used the measures virtually exclusively in this way and usually only for very small sections of phase-space (e.g Crutchfield and Young, 1989; Lopez-Ruiz et al., 1995).

Our results also show that it is possible to use both the Shannon entropy and the statistical complexity to distinguish a complex system from a chaotic and a random one. However, this requires considerable existing knowledge of the systems since a phase-space partition needs to be chosen that adequately resolves a defining feature of complexity. It could be argued that in this case the employment of a complexity measure is probably unnecessary and the systems in question could be classified based on their dynamical properties alone.

Chapter 5

Conclusions

In this chapter we will briefly summarize the results of our study and outline some directions for further research on complexity measures.

Pseudo-randomness and chaos Our discussion in chapter 2 indicates that an epistemological connection between chaos theory and complexity research has been established, which is not based on a conceptual link between these two phenomena but rather induced by incorrect over-use of the term ‘chaos’. We consider chaos to be pseudo-randomness arising in a very restricted class of deterministic, recurrent maps. The systems studied in complexity research are dynamically very different from these, most importantly because they are inter-connected, multi-agent systems. The disorder that can occur in these systems should be correctly described as pseudo-randomness.

Severing the unnecessary linguistic relations between chaos theory and complexity science is not just a pedantic insistence on nomenclature. It has also led to a general tendency to employ chaos theoretical tools in complexity science. The image of ‘complexity at the edge of chaos’ motivated the development of complexity tools that are ‘calibrated’ onto this concept. It is telling that the latest class of complexity tools has been designed and tested on the logistic equation (Crutchfield and Young, 1989; Lopez-Ruiz et al., 1995). As we have argued in chapter 4, these measures are at best awkward in demarcating complex systems.

A tentative definition of complexity Our analysis in chapter 2 shows that the a dynamical definition of complexity can be achieved relatively easily: complex systems are multi-agent systems with directed interactions. However, of these we only want to count the ‘interesting’ ones as complex. For example, while all cellular automata fulfil the dynamical condition, we would ideally like to only classify the class IV ones as complex. Accordingly, we need an additional phenomenological condition. The intuitive notion of complexity is often roughly described as a co-existence of order and disorder. At the moment, the conceptual representation of this idea is mixed up with the (in our opinion: unhelpful) metaphor of the ‘edge of chaos’. In our definition we have tried to avoid this idea and simply required

the co-existence of structures and pseudo-randomness. A second aspect to be considered for a future, successful definition is the crucial influence of the representational possibilities of computer models on the phenomenology of complex systems. This aspect urgently needs further research. In our definition, we only nominally refer to this by requiring a ‘suitable’ representation.

Some ideas for further research on complexity measures Moving away from the unjustified alliance of complexity research and chaos theory, we suggest that further research on complexity measures should focus on two aspects: defining dynamical properties of complexity, namely interconnectedness; and the fact that the phenomenological recognition of complexity is fundamentally anthropocentric. To develop measures quantifying the interconnectedness it will be necessary to consider research like the one by Holland (1975) and learn more about the nature and manifestations of the interactions in complex systems. The phenomenological definition will ultimately have to be based on an anthropocentric notion of what constitutes an interesting pattern. There exists some philosophical and psychological research on this topic that has not been considered for any of the complexity measures so far.

Lloyd (2001) prefaced his list of complexity measures by stating that (p. 7):

“An historical analogue to the problem of measuring complexity is the problem of describing electromagnetism before Maxwell’s equations.”

In a characteristic display of variety, another prominent complexity researcher, Mitchell (2009), described the search for a unified definition of complexity as “waiting for Carnot” (p. 302). Our research seems to indicate that what is needed is not the creative genius of Carnot or Maxwell but is rather the mental precision of Charles Babbage: The man who blithely corrected Lord Tennyson’s poetic estimate of “Every moment dies a man. Every moment one is born.” (*The Vision of Sin*) to the (almost) correct value of $1\frac{1}{16}$ births per death (Babbage and Swade, 2001), would clearly have appreciated the difference between ‘chaos’ and ‘pseudo-randomness’.

Appendix A

MathCad codes for the simulations

```
r := 0..100
t := 0..200

N0,r := 0.1 +  $\frac{(1 - \text{rnd}(2))}{100}$ 

Nt+1,r := Nt,r · e $\frac{r}{20} \left(1 - \frac{N_{t,r}}{0.5}\right)$ 

Nint(t,r) :=  $\frac{0.5}{1 + \frac{\left(1 - \frac{N_0}{0.5}\right) \cdot e^{-\frac{r}{20} \cdot t}}{\frac{N_0}{0.5}}}$ 

WRITEPRN("Population_2.prm") := N
```

Figure A.1: MathCad code for the logistic system.

```

p := 0.2
n := 0..100

x_n := for n ∈ 0..100
    | d ← round(md(100))
    | x_n ← 1 if d ≤ p·100
    | x_n ← 0 otherwise

bl := READPRN("InitialRandom100.prn")
b_{n,0} := bl_n
t := 0..200
b_{n,t} := for t ∈ 0..199
    | for n ∈ 0..99
    | | b_{n+1,t+1} ← |b_{n,t} - 1| if x_n = 1
    | | b_{n+1,t+1} ← b_{n,t} otherwise
    | b_{0,t+1} ← |b_{100,t} - 1| if x_{100} = 1
    | b_{0,t+1} ← b_{100,t} otherwise

w_{n,t} := for t ∈ 0..200
    for n ∈ 0..100
        w_{n,t} ← |b_{n,t} - 1|

WRITEPRN("Black02_100_20.prn") := b
WRITEPRN("White02_100_20.prn") := w
WRITEPRN("X02_100_20.prn") := x

```

Figure A.2: MathCad code for the Kac ring model.

```

n := 0..100

fl := READPRN("Initial_Random_100.prn")
f_{n,0} := fl_n
t := 0..200
f_{n,t} := for t ∈ 1..200
    | row ← f^{(t-1)}
    | minX ← min(row)
    | minXind ← match(minX, row)
    | len ← length(minXind)
    | for n ∈ 0..100
    | | f_{n,t} ← f_{n,t-1}
    | | for m ∈ 0..len - 1
    | | | index ← minXind_m
    | | | f_{index,t} ← md(1)
    | | | f_{index-1,t} ← md(1) if index > 0
    | | | f_{index+1,t} ← md(1) if index < 100
    | | | f_{0,t} ← md(1) if index = 100
    | | | f_{100,t} ← md(1) if index = 0

WRITEPRN("Classic_100_21.prn") := f

```

Figure A.3: MathCad code for the classic Bak-Sneppen model.

```

n := 0..100
k := 2

fn,0 := for n ∈ 0..100
    fn,0 ← rnd(1)

t := 0..200

fn,t := for t ∈ 1..200
    minX ← min(f{t-1})
    minXind ← match(minX, f{t-1})
    len ← length(minXind)
    for n ∈ 0..100
        fn,t ← fn,t-1
    for m ∈ 0..len - 1
        index ← minXindm
        findex,t ← mod(k·findex,t-1, 1)
        findex-1,t ← mod(k·findex-1,t-1, 1) if index > 0
        findex+1,t ← mod(k·findex+1,t-1, 1) if index < 100
        f0,t ← mod(k·f0,t-1, 1) if index = 100
        f100,t ← mod(k·f100,t-1, 1) if index = 0

WRITEPRN("FitnessClassicDet_100_20.prn") := f

```

Figure A.4: MathCad code for the deterministic Bak-Sneppen model.

```

n := 0..100
p := 0.2
fl := READPRN("InitialRandom100.prn")
fn,0 := fln
t := 0..200
fn,t := for t ∈ 1..200
    x ← round(md(100))
    index ← x
    for n ∈ 0..100
        fn,t ← fn,t-1
    if findex,t-1 = 0
        d ← md(100)
        findex,t ← 1 if d ≤ p·100
        findex,t ← 0 otherwise
    if index > 0
        d2 ← md(100)
        findex-1,t ← 1 if d2 ≤ p·100
        findex-1,t ← 0 otherwise
    if index < 100
        d3 ← md(100)
        findex+1,t ← 1 if d3 ≤ p·100
        findex+1,t ← 0 otherwise
    if index = 100
        d4 ← md(100)
        f0,t ← 1 if d4 ≤ p·100
        f0,t ← 0 otherwise
    if index = 0
        d5 ← md(100)
        f100,t ← 1 if d5 ≤ p·100
        f100,t ← 0 otherwise

WRITEPRN("Dis_02_100_20.prn") := f

```

Figure A.5: MathCad code for the discrete Bak-Sneppen model.

Appendix B

Animations

All animations have been uploaded unto CamTools.

KacRun1.wmv: The first fifty iterations of a run from the Kac ring ensemble discussed in chapters 3 and 4.

KacRun2.wmv: The first fifty iterations of a Kac ring starting from an all white (red in the animation) state.

ClassicBS.wmv: The first fifty iterations of a run from the classic Bak-Sneppen ensemble discussed in chapters 3 and 4. Red: Fitness value. Dashed Black: Critical value r_c .

DeterministicBS.wmv: The first fifty iterations of a run from the deterministic Bak-Sneppen ensemble discussed in chapters 3 and 4. Red: Fitness value. Dashed Black: Critical value r_c .

DiscreteBS.wmv: The first fifty iterations of a run from the discrete Bak-Sneppen ensemble discussed in chapters 3 and 4. Red: Zero. Black: One.

LogEq.wmv: The first fifty iterations of a run from the discrete Bak-Sneppen ensemble discussed in chapters 3 and 4.

Bibliography

- W. B. Arthur. The Economy and Complexity. In D. Stein, editor, *Lectures in the Sciences of Complexity*, pages 713–737. Addison-Wesley Longman, Redwood City, 1989.
- C. Babbage and D. Swade. *The difference engine: Charles Babbage and the quest to build the first computer*. Viking Penguin, Ringwood, 2001.
- P. Bak and M. Paczuski. Complexity, contingency, and criticality. *Proceedings of the National Academy of Science USA*, 92:6689–6696, 1995.
- P. Bak and K. Sneppen. Punctuated Equilibrium and Criticality in a Model of Evolution. *Physical Review Letters*, 71:4083–4086, 1993.
- P. Bak, K. Chen, and M. Creutz. Self-organized criticality in the 'Game of Life'. *Nature*, 342:780–782, 1989.
- J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey. On Devaney's Definition of Chaos. *The American Mathematical Monthly*, 99:332–334, 1992.
- D. Bawden. Organized complexity, meaning and understanding. *New Information Perspectives*, 59:307–327, 2007.
- C. Beck. Generalized information and entropy measures in physics. Technical report, 2009.
- C. H. Bennett. On the Nature and Origin of Complexity in Discrete, Homogeneous, Locally-Interacting Systems. *Foundations of Physics*, 16:585–592, 1986.
- C. H. Bennett. How to Define Complexity in Physics, and Why. In N. H. Gregersen, editor, *From Complexity to Life: On the Emergence of Life and Meaning*, pages 34–47. Oxford University Press, Oxford, 2003.
- L. Boltzmann. Weitere Studien ueber das Waermegleichgewicht unter Gasmolekuelen. *Wiener Berichte*, 66:275–370, 1872.
- L. Boltzmann. Bemerkungen ueber einige Probleme der mechanischen Waermetheorie. *Wiener Berichte*, 75:62–100, 1877a.

- L. Boltzmann. Ueber die Beziehung zwischen dem Hauptsatze der mechanischen Waermetheorie und der Wahrscheinlichkeitsrechnung respektive den Saetzen ueber das Waermegleichgewicht. *Wiener Berichte*, 76:373–435, 1877b.
- R. Brown. A brief account of microscopical observations made in the months of june, july and august, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philosophical Magazine*, 4:161–173, 1828.
- D. Byrne. What is Complexity science? Thinking as a Realist about Measurement and Cities and Arguing for Natural History. *Emergence*, 3:61–76, 2001.
- R. Carnap. *Two essays on entropy*. University of California Press, Berkley, 1977.
- G. J. Chaitin. On the length of programs for computing finite binary sequences. *Journal of the Association for Computing Machinery*, 13:547–569, 1966.
- G. J. Chaitin. On the length of programs for computing finite binary sequences: statistical considerations. *Journal of the Association for Computing Machinery*, 16:154–169, 1969.
- D. Chu, R. Strand, and R. Fjelland. Theories of complexity. *Complexity*, 8:19–29, 2003.
- L. O. Chua, S. Yoon, and R. Dogaru. A Nonlinear Dynamics Perspective of Wolfram’s New Kind of Science. Part I: Threshold of Complexity. *International Journal of Bifurcation and Chaos*, 12: 2655–2766, 2002.
- R. Clausius. Ueber verschiedene fuer die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Waermetheorie. *Annalen der Physik und Chemie*, 7:353–400, 1865.
- P. Coveney and R. Highfield. *Frontiers of Complexity: The Search for Order in a Chaotic World*. Faber and Faber, London, 1995.
- J. P. Crutchfield. The calculi of emergence: computation, dynamics and induction. *Physica D*, 74: 11–54, 1994.
- J. P. Crutchfield and K. Young. Inferring Statistical Complexity. *Physical Review Letters*, 63:105–108, 1989.
- P. Davies. Introduction: Towards an emergentist worldview. In N. H. Gregersen, editor, *From Complexity to Life: On the Emergence of Life and Meaning*, pages 3–19. Oxford University Press, Oxford, 2003.
- P. de Los Rios, A. Valleriani, and J. L. Vega. Self-organized criticality driven by deterministic rules. *Physical Review E*, 56:4876–4879, 1997.
- R. L. Devaney. *An Introduction To Chaotic Dynamical Systems*. Addison Wesley, Redwood City, 1989.

- J. R. Dorfman. *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*. Cambridge University Press, Cambridge, 1999.
- B. Edmonds. *Syntactic measures of complexity*. PhD thesis, University of Manchester, 1999.
- P. Ehrenfest and T. Ehrenfest. *The conceptual foundations of the statistical approach in mechanics*. Cornell University Press, Ithaca, 1959.
- M. J. Feigenbaum. Quantitative Universality for a Class of Nonlinear Transformations. *Journal of Statistical Physics*, 19:25–52, 1977.
- D. P. Feldman and J. P. Crutchfield. Statistical Measures of Complexity, Why? *Physics Letters A*, 238: 244–252, 1998.
- R. Frigg. Self-organized criticality - what it is and what it isn't. *Studies of the History and Philosophy of Science*, 34:613–632, 2003.
- R. Frigg and C. Werndl. Entropy - A guide for the perplexed. preprint, available at philsci-archive.pitt.edu/8592/, 2011.
- N. Ganguly, B. K. Sikdar, A. Deutsch, G. Canright, and P. P. Chaudhuri. A Survey on Cellular Automata. Technical report, 2003.
- M. Gardner. Mathematical Games: The fantastic combinations of john conway's new solitaire game 'Life'. *Scientific American*, 223:120–123, 1970.
- M. Gell-Mann. *The Quark and the Jaguar: Adventures in the Simple and the Complex*. W. H. Freeman, New York, 1994.
- M. Gell-Mann. What is complexity? *Complexity*, 1:1–9, 1995.
- M. Gell-Mann and S. Lloyd. Effective complexity. In M. Gell-Mann and C. Tsallis, editors, *Nonextensive Entropy: Interdisciplinary Applications*, pages 387–398. Oxford University Press, Oxford, 2004.
- C. Gershenson. *Five Questions on Complexity*. Automatic Press, Copenhagen, 2008.
- A. J. Gillet. *Phase transitions in Bak-Sneppen avalanches and in a continuum percolation model*. PhD thesis, Vrije Universiteit Amsterdam, 2007.
- J. Gleick. *Chaos: Making a New Science*. Viking Penguin, New York, 1987.
- D. E. Goldberg and J. Holland. Genetic Algorithms and Machine Learning. *Machine Learning*, 3:95–99, 1988.
- G. A. Gottwald and M. Oliver. Boltzmann's Dilemma: An Introduction to Statistical Mechanics via the Kac Ring. *SIAM Review*, 51:613–635, 2009.

- P. Grassberger. The Bak-Sneppen model for punctuated evolution. *Physics Letters A*, 200:277–282, 1995.
- P. Grassberger and I. Procaccia. Characterization of Strange Attractors. *Physical Review Letters*, 50: 346–349, 1983.
- N. H. Gregersen. *From Complexity to Life: On the Emergence of Life and Meaning*. Oxford University Press, Oxford, 2003.
- R. C. Hilborn. *Chaos and Nonlinear Dynamics*. Oxford University Press, Oxford, 2002.
- J. H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, 1975.
- J. H. Holland. Complex Adaptive Systems. *Daedalus*, 121:17–30, 1992.
- J. Horgan. From complexity to perplexity. *Scientific American*, 272:104–110, 1995.
- R. L. Hotz. A Study in Complexity. *Technological review*, pages 23–29, October 1997.
- E. T. Jaynes. Information Theory and Statistical Mechanics. *The Physical Review*, 106:620–630, 1957a.
- E. T. Jaynes. Information Theory and Statistical Mechanics. II. *The Physical Review*, 108:171–190, 1957b.
- M. Kac. Some remarks on the use of probability on classical statistical mechanics. *Bulletin of the Royal Belgium Academy of Science*, 53:356–361, 1956.
- S. Kauffman. *At Home In the Universe: The Search for Laws of Order and Complexity*. Oxford University Press, Oxford, 1995.
- S. H. Kellert. *In the Wake of Chaos*. The University of Chicago Press, Chicago, 1993.
- A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1:1–7, 1965.
- J. Ladyman, J. Lambert, and K. Wisener. What is a complex system? preprint, available at philsci-archive.pitt.edu/8496/, 2011.
- C. G. Langton. Computation at the Edge of Chaos: Phase Transitions and Emergent Computation. *Physica D*, 42:12–37, 1990.
- J. S. Lansing. Complex Adaptive Systems. *Annual Review of Anthropology*, 32:183–204, 2003.
- S. Lloyd. Measures of complexity: A nonexhaustive list. *IEEE Control Systems Magazine*, 21:7–8, 2001.
- S. Lloyd and H. Pagels. Complexity as Thermodynamic Depth. *Annals of Physics*, 188:186–213, 1988.

- R. Lopez-Ruiz, H. L. Mancini, and X. Calbet. A statistical measure of complexity. *Physics Letters A*, 209:321–326, 1995.
- M. T. Martin, A. Plastino, and O. A. Rosso. Statistical complexity and disequilibrium. *Physics Letters A*, 311:126–132, 2003.
- R. M. May. Biological Populations with Non-overlapping generations: Stable points, Stable Cycles, and Chaos. *Science*, 15:645–647, 1974.
- R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976a.
- R. M. May. Bifurcations and Dynamic Complexity in Simple Ecological Models. *The American Naturalist*, 110:573–599, 1976b.
- B. McKelvey. What is Complexity Science? Is it Really Order-Creation Science? *Emergence*, 3:137–57, 2001.
- B. L. McNaughton. The Neurobiology of Spatial Computation and Learning. In D. Stein, editor, *Lectures in the Sciences of Complexity*, pages 389–435. Addison-Wesley Longman, Redwood City, 1989.
- R. Meester and D. Znamenski. Non-Triviality of a Discrete Bak-Sneppen Evolution Model. *Journal of Statistical Physics*, 109:987–1004, 2002.
- R. V. Mendes. Deterministic Bak-Sneppen model: Lyapunov spectrum and avalanches as return times. *Physics D*, 214:182–186, 2006.
- M. Mitchell. *Complexity: A Guided Tour*. Oxford University Press, Oxford, 2009.
- M. Mitchell, P. T. Hraber, and J. P. Crutchfield. Revisiting the Edge of Chaos: Evolving Cellular Automata to Perform Computations. *Complex Systems*, 7:89–130, 1993.
- M. Mitchell, J. P. Crutchfield, and R. Das. Evolving cellular automata with genetic algorithms: A review of recent work. In *in Proceedings of EvCa'96*, pages 1–14. Russian Academy of Science, 1996.
- V. Nannen. A Short Introduction to Kolmogorov Complexity. *Computing Research Repository*, 2010.
- D. V. Newman. Emergence and Strange Attractors. *Philosophy of Science*, 63:245–261, 1996.
- Oxford Dictionaries. Complexity, 2011. URL <http://oxforddictionaries.com/definition/complexity>. Accessed 27/05/2011.
- N. H. Packard. Adaptation towards the edge of chaos. In J. A. S. Kelso, A. J. Mandell, and M. F. Shlesinger, editors, *Dynamic Patterns in Complex Systems*, pages 293–301. World Scientific, Singapore, 1988.

- N. H. Packard and S. Wolfram. Two dimensional cellular automata. *Journal of Statistical Physics*, 38: 901, 1985.
- S. E. Phelan. What is Complexity Science, Really? *Emergence*, 3:120–136, 2001.
- I. Prigogine. *From Being to Becoming*. W. H. Freeman, Oxford, 1980.
- I. Prigogine and I. Stengers. *Order out of chaos*. Bantam Books, New York, 1984.
- K. Richardson and P. Cilliers. What is Complexity Science? A View from Different Directions. *Emergence*, 3:5–24, 2001.
- R. Rosen. Complexity as a system property. *International Journal of General Systems*, 3:227–232, 1977.
- Santa Fe Institute. History of the Santa Fe Institute, 2011. URL www.santafe.edu/about/history/. Accessed 27/05/2011.
- F. Schaezting. *The Swarm*. Regan Books, New York, 2006.
- J. P. Sethna. *Statistical Mechanics: Entropy, Order Parameters, and Complexity*. Oxford University Press, Oxford, 2006.
- C. E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27: 379–423, 1948.
- M. Shermer. *Why People Believe Weird Things: Pseudo-science, Superstition and Bogus Notions of Our Time*. W. H. Freeman, New York, 2002.
- P. T. Shiner, M. Davison, and P. T. Landsberg. Simple measure for complexity. *Physical Review E*, 59: 1459–1464, 1999.
- P. Smith. *Explaining Chaos*. Cambridge University Press, Cambridge, 1998.
- R. Solomonoff. A formal theory of inductive inference, part 1 and part 2. *Information and Control*, 7: 1–22;224–254, 1964.
- R. K. Standish. Concept and Definition of Complexity. In A. Yang and Y. Shan, editors, *Intelligent Complex Adaptive Systems*, pages 105–124. IGI Global, Hershey, 2008.
- M. Strevens. *Bigger than Chaos*. Harvard University Press, Cambridge, 2003.
- A. M. Turing. On computable numbers with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, Series 2*, 42:230–265, 1936.
- J. Uffink. Compendium of the foundations of classical statistical physics. In J. Butterfield and J. Earman, editors, *Philosophy of Physics*, pages 923–1074. Elsevier, New York, 2007.
- J. Uffink. Boltzmann’s Work in Statistical Physics, 2008. URL plato.stanford.edu/archives/win2008/entries/statphys-Boltzmann. Accessed 27/05/2011.

- J. Von Neumann and A. W. Burk (Ed.). *Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana, 1963.
- M Mitchell Waldrop. *Complexity: The emerging science at the edge of order and chaos*. Penguin Books, London, 1992.
- W. Weaver. Science and Complexity. *The American Scientist*, 56:536–547, 1948.
- E. W. Weisstein. Rule 30, 2011. URL <http://mathworld.wolfram.com/Rule30.html>. Accessed 30/05/2011.
- C. Werndl. Are deterministic and indeterministic descriptions observationally equivalent. *Studies of History and Philosophy of Modern Physics*, 40:232–242, 2009.
- S. Wolfram. Universality and Complexity in Cellular Automata. *Physica D*, 10:1–35, 1984.
- S. Wolfram. *A New Kind of Science*. Wolfram Media, Champaign, 2002.
- W. H. Zurek. Algorithmic Information Content, Church-Turing Thesis, Physical Entropy and Maxwell’s Demon. In W. H. Zurek, editor, *Complexity, Entropy and the Physics of Information*, pages 73–90. Addison-Wesley Longman, Redwood City, 1990.