Neutral Emergence: a proposal *

Andrew Weeks, Susan Stepney, Fiona Polack
Department of Computer Science, University of York, UK

Abstract

In this position paper, we introduce the concept of *neutral emergence* (defined by analogy to an information theoretic view of neutral evolution), and discuss how it might be used in the engineering of *robust* emergent systems.

1 Introduction

We are interested in engineering emergent systems. That is, given a specification of a property, how can we implement a system such that this property emerges? Furthermore, how can we make this property robust, and how can we argue that the system is suitably implemented (for example, for certification purposes)?

In this position paper, we introduce the concept of *neutral emergence* (defined by analogy to an information theoretic view of neutral evolution), and discuss how it might be a potential component of such an engineering process.

Section 2 summarises background material on neutral evolution, and its formulation in information theoretic terms. Section 3 summarises background material on emergence. In section 4 we cast emergence in information theoretic terms, by analogy to information theoretic evolution. In section 5 we introduce the concept of *neutral emergence*, show how it is related to system robustness, and describe our plans to investigate this further.

*For discussion at the Symposium on Complex Systems Engineering, RAND Corporation, Santa Monica, CA, USA, 11–12 January 2007

2 Background: evolution

2.1 Neutral evolution

The mapping from genotype (DNA) to phenotype (organism) is complex, and there is significant redundancy in both. Different genotypes can map to the same phenotype; for example, different codons (DNA nucleotide triplets) can code for the same amino acid. Hence the genotype can change (a nucleotide can mutate) without changing the phenotype. Similarly, the same genotype can result in different phenotypes, due to different environmental conditions during development.

Neutral evolution is a well recognised phenomenon in evolutionary biology [Stearns & Hoekstra 2000, Ridley 2004]. It involves a change in the genotype without a significant change in the fitness of the phenotype; that is, it involves a change that is selectively neutral. Such changes allow the phenotypic population to explore their fitness landscape, by drifting along contours of equal fitness, to regions where they may subsequently find fitter solutions not directly accessible from their original position in the landscape.

2.2 Information theoretic evolution

[Adami 1998, Adami & Cerf 2000, Adami 2002] gives a description of evolution in information-theoretic terms, by considering the information in the genome *in the context of its environment*.

[Adami 1998] classifies a (digital) genome into 'hot' and 'cold' bits. The cold bits are ones strongly conserved in the population, whereas the hot bits vary between population members, and are therefore (presumably) not significantly contributing to the fitness. So, in general, mutations of hot bits are *neutral*, but occasionally might be beneficial, at which point they become 'frozen in' to the population, and conserved.

These fitter organisms are exploiting their environment better, and so must contain more information about the environment [Adami 1998, p115]. In information theoretic evolution, the genome is seen as some kind of representation of the environment: the cold bits are strongly correlated with the environment, whereas the hot bits are not. The more correlated bits there are, the higher the *mutual information* between the organism's genome and the environment:

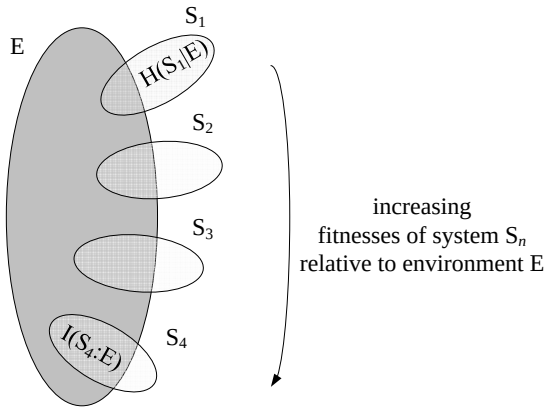

Figure 1: Evolution increasing the mutual information between the genome S and the environment E. (Adapted from [Adami 1998, fig.5.7].)

$$I(S : E) = H(S) - H(S|E) \tag{1}$$

The *mutual information*, or correlation, between the system S and its environment E, $I(S : E)$, is the *entropy* of the system, $H(S)$, less the *conditional entropy* of the system in the context of the environment, $H(S|E)$. This conditional entropy $H(S|E)$ can be thought of as the amount of information in the system that cannot be explained by (correlations with) its environment.

Evolution (increasing fitness) is then *increasing mutual information*: increasing the shared information, or correlations, between the genome and the environment. So "*natural selection can be viewed as a* filter *... that lets information flow into the genome, but prevents it from flowing out*" [Adami 2002]. See figure 1. (For simplicity, the figure shows the total information in S as not changing, hence the conditional information decreases as the mutual information increases. This is not necessarily the case: a neutral evolutionary step could change the amount of conditional information $H(S|E)$ by increasing the size of the uncorrelated part of the genome [Adami 2002]. This could happen, for example, by mutating a redundant part of the correlated genome such as could have resulted from a gene duplication event.)

[Adami 1998] allows bits to 'freeze' individually, with monotonic increase in fitness. In reality, things are more complicated, as bits are correlated within a genome [Adami 2002]. Evolutionary search is a complex combinatorial problem, so when some bits go cold it is necessary for other bits to

become hot again: there is no guarantee that the mutual information will increase monotonically. This can affect evolutionary trajectories, as it may not be possible 'to get there from here' by the moves permitted by evolutionary operators [Kauffman 1993]. However, this does not undermine the principle of the information-theoretic approach.

2.3 Neutral evolution in information theoretic terms

The mutual information, I, is independent of the hot bits in S; it depends only on the cold bits correlated with the environment E. These random hot bits can change without affecting the mutual, or shared, information. Hence, in this context, neutral evolution is a change in the genome that does not affect the mutual information of the genome and its environment, $I(S : E)$. This means that S is *robust* to these kind of genomic changes.

As noted earlier, the conditional information $H(S|E)$ may increase without decreasing the fitness if, for example, a mutation occurs in a redundant part of the genome. [Soule 2003] discusses genome growth as a strategy for increasing robustness (to perturbations from the genetic operators).

A similar argument means that parts of the environment E that are not correlated with S can also change neutrally (from the point of view of S): can change without the organism "noticing". So S is *robust* to these kind of environmental changes, too.

When using an artificial fitness function in an evolutionary algorithm, it is important to take these ideas into account to get a robust solution. [Branke 1998] notes that "*this means that not only the solution should be good, but also that the (phenotypic) neighbourhood of the solution should have a high average quality. Looking at the fitness landscape, a solution on a high plateau should be preferred over a solution on a thin peak: if the environment changes slightly . . . the solution on the plateau will yield much better expected quality than the solution on the peak.*" This can be achieved by adding noise to the fitness function, thereby increasing $H(E|S)$.

2.4 Information theoretic measures

As [Adami 1998, Adami 2002] points out, what is important is the mutual information between the system and its environment: the amount of information stored in the system about its environment. [Adami & Cerf 2000]

dub this mutual information the *physical complexity* (a relative measure; as opposed to the mathematical, or intrinsic, complexity).

Calculating this requires the use of some information theoretic measure (as seen in equation 1). In this paper, we are deliberately vague about the precise form of measure to be used, as the details depend on the particular measure, but the concepts are not so sensitive to it. Some measures are summarised in appendix A.

3 Background: Emergence

3.1 Definitions

[Stepney *et al.* 2006] provide a review of opinion on emergence, and sum up its characteristics as follows.

- The whole (system) is greater than, and different from, the sum of its parts (in terms of behaviour and other qualities) [Aristotle 350 BC, Anderson 1972].

- Emergence can, in various senses, be equated to novelty [Crutchfield 1994, Ronald *et al.* 1999, Bickhard & Campbell 2000] (but not to 'surprise'; see later).

- A process view, in which particles are merely expressions of stable processes, is more useful than a substance view [Bickhard & Campbell 2000, Campbell & Bickhard 2001, Abbott 2006], especially when considering emergence in non-equilibrium dynamical systems.

- Levels are essential: emergence occurs at a higher level [Emmeche *et al.* 1997, Hordijk *et al.* 1998, Bickhard & Campbell 2000] or longer time scale than that on which basic processes occur [Campbell 1974, Burns *et al.* 2005].

- Emergence is a characteristic that is expressed at the high level, caused by the low level [Bedau 1997, Abbott 2006], yet emergent systems may exhibit a form of 'downward' causality, where the high level emergence influences subsequent low-level activity [Campbell 1974, O'Conner 1994, Faith 1998, Bickhard & Campbell 2000, Campbell & Bickhard 2001].

- Emergence is related to the attractors of dynamical systems [Kauffman 1993, Newman 1996, Goldstein 1999].

[Bar-Yam 2004] argues that *strong* emergence is a consequence of global constraints on the system (as opposed to constraints on its components). As a subsystem cannot be understood in isolation from the system, so a system cannot be understood in isolation from its environment.

Our starting point is (an adaptation of) [Ronald *et al.* 1999]'s definition of emergence. They say "*The language of design L_1 and the language of observation L_2 are distinct, and the causal link between the elementary interactions programmed in L_1 and the behaviors observed in L_2 is* non-obvious *to the observer—who therefore experiences surprise.*" In [Stepney *et al.* 2006] we explain why we do not like the use of 'surprise' as a criterion for emergence: it is subjective, and wears out with repetition. However, we do follow [Ronald *et al.* 1999]'s use of two languages of description, which here we call L for the 'microscopic' *local* level of implementation substrate, and S for the 'macroscopic' global level of the model or *specification* (we are interested both in observing and modelling natural emergent systems, and in specifying and implementing engineered ones).

3.2 Existence of 'natural' levels

[Israeli & Goldenfeld 2006] note that there is an emergent natural length scale for CA coarse graining (see appendix B). Other such processes also have an associated natural scale. For example, the process of "reconstructing the attractor" from time-lagged observations [Takens 1981] is a form of coarse-graining that does not lose the underlying dynamics, and the correct time lag to use can be found using mutual information [Ray 2004].

Hence emergence is not an arbitrary feature, dependent merely on the chosen coarse graining or level of observation. There are some levels that are 'better' than others (for example, coarse grainings that better describing the underlying dynamics), which form a natural level for discovering, or designing, emergent properties.

4 Emergence in information theoretic terms

By direct analogy to the information theoretic description of evolution, we can define an information theoretic quantity of emergence:

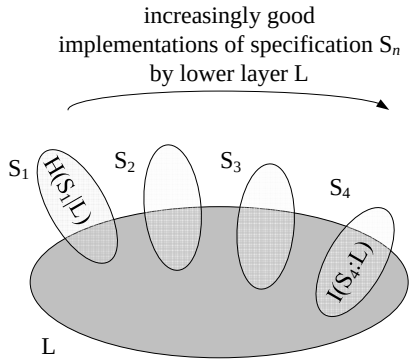

Figure 2: Modelling, or incremental system development, as increasing the mutual information between the system specification S and the implementation L.

$$I(S:L) = H(S) - H(S|L) \tag{2}$$

The amount of emergence $I(S:L)$ is the mutual information, or correlation, between the specified system S and its implementation in substrate level L. This is the entropy of the specified system, $H(S)$, less the conditional entropy of the system in the context of the implementation substrate, $H(S|L)$. This conditional entropy $H(S|L)$ can be thought of as the amount of information in the system specification that has not been captured by (correlations with) its implementation.

Modelling, or incremental system development, can be viewed as increasing mutual information: increasing the shared information, or correlations, between the system specification and its implementation (figure 2).

If one were trying to explain (model) an observed system S in terms of L, the conditional information $H(S|L)$ might be considered to be 'surprise': behaviours or properties of S not explained by L. (It might just be noise.) If one were trying to implement a specified system S in an implementation substrate L, this conditional information is the part of the specification that has yet to be captured by the proposed implementation L: more development work is required.

The conditional information $H(L|S)$ (the information in the implementation not correlated with the system specification) might be considered as 'surprising' properties of L, properties unnecessary for the realisation of S. Of course, these additional properties might well be *invisible* if the system

is viewed through high level observation 'glasses' [Clark *et al.* 2005] that see only system-level properties (multiple distinct microstates nevertheless resulting in the same observed macrostate). We see (later) that these extra properties can nevertheless be exploited to get *robust* implementations of S.

This suggests an approach to incremental system development: use the mutual information I as a fitness function to search for good models (system descriptions) S of an existing L (such as a given CA or agent system), or to search for good low level implementations L (such as the required rules and states for a CA or agent system) of a system specification S.

5 Neutral Emergence

5.1 Definition

An emergent property exhibits *neutral emergence* when a change in the microstate L does not change the macrostate S, or vice versa. It is a change that does not affect the mutual information $I(S : L)$.

5.2 Robustness

So the system S is robust to changes in either itself, or in L, that do not change the mutual information. In particular, it can be robust to many changes in its implementation, including, possibly, the effect of errors.

It is often stated that emergent systems (often modelled on natural processes) exhibit robustness: here we see why (and where) this may be the case. The excess information in L (a large $H(L|S)$) is *necessary* for emergent systems to be robust in this manner.

As argued earlier, an engineering development process can be seen as implementing specification S by finding an L with a high mutual information $I(S : L)$. Here we see that, at the same time, the process can also seek to maximise robustness, by searching for a system that is insensitive to (uncorrelated with) certain failure modes or other possible changes in L. If a system were stressed during development (exposed to a range of stresses and implementation errors), its implementation could be encouraged towards regions that are insensitive (robust) to such events. (Compare this to the development of formally proven systems: they do not guarantee any level of performance with even the smallest change.) By analogy to evolutionary

fitness landscapes, we want to find systems that lie in gently sloping plains and plateaux, rather than on narrow peaks or steep cliffs.

5.3 Speculations on 'the edge of emergence'

Experiments with CAs [Langton 1991] and Random Boolean Networks [Kauffman 1993, Kauffman 1995] indicate that a phase transition takes place between a 'solid' ordered regime and a 'gaseous' chaotic state, and that the narrow 'liquid' region perched on the boundary between these regimes is capable of supporting the most complex behaviour, including universal computation. (Note that the details of Langton's results have been questioned by [Mitchell *et al.* 1994].)

Compare this with our mutual information measure of emergence: a system exhibits minimal emergence when everything is a surprise (zero mutual information). Clearly a model like this that knows nothing about what it is modelling is useless, but equally (as argued above) some degree of surprise (some conditional information, or less than maximal mutual information) in the system may prove advantageous. Thus there should be a level of emergence with the maximum utility, a position at which the model has freedom to explore but is held within a constrained region of the search space. While it is tempting to label this the 'emergence liquid region', substantive experimental data should obtained before making such a claim.

That said, consideration of the analogy raises some interesting questions. Does a phase transition occur between the 'solid' (system specification or model effectively duplicating the implementation) and 'gaseous' (little in common) regimes? Is the region a narrow one? If so, is this why creating effective emergent systems manually has proved difficult? Will automatic emergent systems evolve towards this liquid boundary as Kauffman suggests life did?

5.4 Next Steps

Our next steps are to investigate the properties of neutral emergence and associated robustness in several classes of systems: Game of Life, flocking, and simple substitution ciphers.

We are starting from the CA coarse-graining ideas of [Israeli & Goldenfeld 2006] (appendix B), recasting them in the information theoretic emergent framework sketched above, and applying them to make predictions about

glider behaviour in Conway's Game of Life CA [Berlekamp *et al.* 1982]. In particular, we are investigating coarse graining defined over only some CA states, thereby allowing the information content of those states to be exploited, in addition to the information content of the CA rules. (Compare the case in physical emergent systems, where the emergent properties occur only over some restricted set of all possible states, such as a restricted temperature range.)

We will then move to continuous systems, and apply the ideas to design robust emergent flocking behaviours.

We are also investigating various cryptographic attacks on simple transpositions ciphers (such as described in [Matthews 1993, Russell *et al.* 2003]), in order to explore the use of ideas of conditional entropy to maximise robustness: in this case, robustness of the cracking mechanism to changes in the fitness function of bigram and trigram frequency distributions.

Eventually, we want to examine emergence in the context of an environment (combining some of the ideas of evolution and emergence mentioned herein), both for defining emergent properties, and for achieving robustness of that emergence.

6 Conclusions

We have cast emergence in information theoretic terms, by direct analogy to evolutionary processes, and shown how this can be used to derive a fitness function for developing models of emergent systems, or emergent implementations of system specifications. We have introduced the concept of neutral emergence, and shown its relationship to robust emergence.

Our eventual aim is to incorporate neutral emergence as a component in a robust emergent engineering process.

Acknowledgements

Andrew Weeks is funded by a Microsoft Research European PhD Scholarship.

References

[Abbott 2006]
Russ Abbott. Emergence explained: getting epiphenomena to do real work. February 2006. arXiv:cs/0602045.

[Abraham 1987]
Ralph H. Abraham. Dynamics and self-organization. In F. E. Yates, A. Garfinkel, D. Walter, and G. Yates, editors, *Self-organizing Systems: The Emergence of Order*, pages 599–613. Plenum, 1987.

[Adami & Cerf 2000]
Christoph Adami and N. J. Cerf. Physical complexity of symbolic sequences. *Physica D*, 137:62–69, 2000.

[Adami 1998]
Christoph Adami. *Introduction to Artificial Life*. Springer, 1998.

[Adami 2002]
Christoph Adami. What is complexity? *BioEssays*, 24:1085–1094, 2002.

[Anderson 1972]
Philip W. Anderson. More is different. *Science*, 177(4047):393–396, 1972.

[Aristotle 350 BC]
Aristotle. *Metaphysics*, volume book H (VIII). 350 BC. Translation from W. D. Ross, *Aristotle's metaphysics*, 2 vols, Oxford University Press, 1924.

[Badii & Politi 1997]
Remo Badii and Antonio Politi. *Complexity: hierarchical structures and scaling in physics*. Cambridge University Press, 1997.

[Bar-Yam 2004]
Yaneer Bar-Yam. A mathematical theory of strong emergence using multiscale variety. *Complexity*, 9(6):15–24, 2004.

[Bedau 1997]
Mark A. Bedau. Weak emergence. In J. Tomberlin, editor, *Philosophical Perspectives 11: Mind, Causation, and World*, pages 375–399. Blackwell, 1997.

[Berlekamp *et al.* 1982]
E. R. Berlekamp, J. H. Conway, and R. K. Guy. *Winning Ways for your Mathematical Plays*. Academic Press, 1982.

[Bickhard & Campbell 2000]
Mark H. Bickhard and D. T. Campbell. Emergence. In P. B. Andersen, C. Emmeche, N. O. Finnemann, and P. V. Christiansen, editors, *Downward Causation*, pages 322–348. Aarhus University Press, 2000.

[Branke 1998]
Jürgen Branke. Creating robust solutions by means of evolutionary algorithms. In [Eiben *et al.* 1998], pages 119–128.

[Burns *et al.* 2005]
A. Burns, I. J. Hayes, G. Baxter, and C. J. Fidge. Modelling temporal behaviour in complex socio-technical systems. Technical Report YCS-2005-390, Department of Computer Science, University of York, 2005.

[Campbell & Bickhard 2001]

Richard J. Campbell and Mark H. Bickhard. Physicalism, emergence and downward causation. 2001. www.lehigh.edu/~mhb0/physicalemergence.pdf.

[Campbell 1974]
Donald T. Campbell. 'Downward Causation' in hierarchically organised biological systems. In Francisco Jose Ayala and Theodosius Dobzhansky, editors, *Studies in the Philosophy of Biology: reduction and related problems*, chapter 11, pages 179–186. Macmillan, 1974.

[Clark *et al.* 2005]
John A. Clark, Susan Stepney, and Howard Chivers. Breaking the model: finalisation and a taxonomy of security attacks. In *REFINE 2005 Workshop, Guildford, UK, April 2005*, volume 137(2) of *ENTCS*, pages 225–242. Elsevier, 2005.

[Crutchfield 1992]
James P. Crutchfield. Semantics and thermodynamics. In Martin Casdagli and Stephen Eubank, editors, *Nonlinear Modeling and Forecasting*, pages 317–359. Addison Wesley, 1992.

[Crutchfield 1994]
J. P. Crutchfield. The calculi of emergence: Computation, dynamics, and induction. In *Physica D special issue on the Proceedings of the Oji International Seminar*, 1994.

[Eiben *et al.* 1998]
A. E. Eiben, T. Bäck, M. Schoenauer, and H.-P. Schwefel, editors. *Parallel Problem Solving from Nature — PPSN V*, volume 1498 of *LNCS*. Springer, 1998.

[Emmeche *et al.* 1997]
C. Emmeche, S. Køppe, and F. Stjernfelt. Explaining emergence: towards an ontology of levels. *Journal for General Philosophy of Science*, 28:83–119, 1997.

[Faith 1998]
J. Faith. Why gliders don't exist: Anti-reductionism and emergence. In C. Adami, R.K. Belew, H. Kitano, and C. Taylor, editors, *Artificial Life VI: Proceedings of the Sixth International Conference on Artificial Life*, pages 389–392, 1998.

[Goldstein 1999]
Jeffrey Goldstein. Emergence as a construct: History and issues. *Emergence*, 1(1):49–72, 1999.

[Hordijk *et al.* 1998]
W. Hordijk, J. P. Crutchfield, and M. Mitchell. Mechanisms of emergent computation in cellular automata. In [Eiben *et al.* 1998], pages 613–622.

[Israeli & Goldenfeld 2006]
Navot Israeli and Nigel Goldenfeld. Coarse-graining of cellular automata, emergence, and the predictability of complex systems. *Phys. Rev. E*, 73:026203, 2006.

[Kauffman 1993]
Stuart A. Kauffman. *The Origins of Order: self-organization and selection in evolution*. Oxford University Press, 1993.

[Kauffman 1995]
Stuart A. Kauffman. *At Home in the Universe*. Oxford University Press, 1995.

[Kolmogorov 1965]
A. N. Kolmogorov. Three approaches to the quantitive definition of information. *Problems of Information Transmission*, 1:1–7, 1965.

[Langton 1991]
C. G. Langton. *Computation at the Edge of Chaos: Phase-Transitions and Emergent Computation*. PhD thesis, University of Michigan, 1991.

[Matthews 1993]
Robert A. J. Matthews. The use of genetic algorithms in cryptanalysis. *Cryptologia*, 17(2):187–201, 1993.

[Mitchell *et al.* 1994]
Melanie Mitchell, James P. Crutchfield, and Peter T. Hraber. Dynamics, computation, and the 'edge of chaos': A re-examination. In G. A. Cowan, D. Pines, and D. Meltzer, editors, *Complexity: Metaphors, Models, and Reality*, pages 497–513. Addison Wesley, 1994.

[Newman 1996]
David V. Newman. Emergence and strange attractors. *Philosophy of Science*, 63(2):245–261, 1996.

[O'Conner 1994]
T. O'Conner. Emergent properties. *Americal Philosophical Quarterly*, 31:91–104, 1994.

[Ray 2004]
Asok Ray. Symbolic dynamic analysis of complex systems for anomoly detection. *Signal Processing*, 84:1115–1130, 2004.

[Ridley 2004]
M. Ridley. *Evolution*. Blackwell, 2004.

[Ronald *et al.* 1999]
E. M. A. Ronald, M. Sipper, and M. S. Capcarrère. Testing for emergence in artificial life. In D. Floreano, J.-D. Nicoud, and F. Mondada, editors, *Advances in Artificial Life: 5th European Conference*, volume 1674 of *LNCS*, pages 13–20. Springer, 1999.

[Russell *et al.* 2003]
Matthew D. Russell, John A. Clark, and Susan Stepney. Making the most of two heuristics: Breaking transposition ciphers with ants. In *CEC 2003: International Conference on Evolutionary Computation, Canberra, Australia, December 2003*, pages 2653–2658. IEEE, 2003.

[Ryan 2006]
Alex J. Ryan. Emergence is coupled to scope, not level. 2006. arXiv:nlin.AO/0609011.

[Schmidhuber 1997]
J. Schmidhuber. Discovering neural nets with low Kolmogorov complexity and high generalization capability. *Neural Networks*, 10(5):857–873, 1997.

[Shalizi & Moore 2006]
Cosma Rohilla Shalizi and Cristopher Moore. What is a macrostate? subjective observations and objective dynamics. 2006. arXiv:cond-mat/0303625.

[Shannon 1948]
Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423,623–656, July, October 1948.

[Soule 2003]
Terence Soule. Operator choice and the evolution of robust solutions. In Rick L. Riolo

and Bill Worzel, editors, *Genetic Programming Theory and Practice*, chapter 16, pages 257–270. Kluwer, 2003.

[Stearns & Hoekstra 2000]
S. C. Stearns and R. F. Hoekstra. *Evolution: an introduction*. Oxford University Press, 2000.

[Stepney *et al.* 2006]
Susan Stepney, Fiona Polack, and Heather Turner. Engineering emergence. In *CEC 2006: 11th IEEE International Conference on Engineering of Complex Computer Systems, Stanford, CA, USA, August 2006*. IEEE, 2006.

[Takens 1981]
F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L.-S. Young, editors, *Dynamical Systems and Turbulence*, volume 898 of *LNM*, pages 230–242. Springer, 1981.

A Information theoretic measures

The information (Shannon) entropy [Shannon 1948] is defined over an *ensemble* (representative collection, or population) of systems, in terms of probability p_i of being in each of the possible system states $i \in \mathcal{I}$, the probabilities defined over the ensemble.

$$H(S) = -\sum_{i \in \mathcal{I}} p_i \log_2 p_i \tag{3}$$

[Adami 1998]'s concept of 'hot' and 'cold' bits are defined over an ensemble.

Kolmogorov (or Kolmogorov-Chaitin) complexity [Kolmogorov 1965] can be used to capture the complexity, or compressibility, of a single string (as opposed to an ensemble of strings), in terms of the shortest program that can generate that string. Averaging the Kolmogorov complexity of an ensemble of strings (in the limit of infinitely long strings) yields the information entropy [Adami & Cerf 2000].

These conventional information measures assign minimal values to completely ordered systems, and maximum values to completely random systems. [Crutchfield 1992] defines *statistical complexity*, which assigns a low complexity value to both complete order and complete randomness. This measure is defined in terms of the smallest finite probabilistic automaton that can simulate the statistical properties of the system: "*There will be somewhere in the Chomsky hierarchy an optimal representation which is finitely expressed in the language of the least powerful class*" [Crutchfield 1992]. So complete randomness, for example, is simulated by a trivial 'coin tossing' automaton.

There is a relationship here to ideas from Machine Learning, in the trade-off between classifier size and complexity, and their generalisation capabilities (see, for example, [Schmidhuber 1997]).

B Coarse graining

B.1 Dynamical systems view

Insight arises when emergence is considered as a characteristic of dynamical systems. In a dynamical systems view, everything is process, or a motion towards or on an attractor. However, when the system is monitored at a suitable (suitably large) timescale, it is observed to behave like a particle, because the "slow" observer sees only the average of a series of states [Abraham 1987, Goldstein 1999], not the detailed motion. [Stepney *et al.* 2006] note that, "*if everything is process, then things can be independent of the details of the underlying substrate (composition of the underlying levels): the same processes could be supported by different substrates. This gives us some hope of a science of emergence that is (relatively) independent of the details of specific implementations, and depends just on their dynamical properties (attractor structures).*"

Considerations of emergence in terms of mutual information also lead to independence from certain details of the underlying substrate (section 5.2).

B.2 Symbolic dynamics

Symbolic dynamics is the discipline of recasting a continuous (space and time) dynamical system into a discrete (space and time) one. The phase space of the continuous system is partitioned into a finite number of sets, each labelled with a unique element from a finite alphabet. The system is observed at discretised time intervals, and the label of the set it occupies noted. This process results in a sequence of symbols, and the dynamics of sequences can be analysed. (See, for example, [Badii & Politi 1997, ch.4].) Under suitable partitionings, the dynamics of the sequences is a good indicator of the dynamics of the underlying system.

Symbolic dynamics is a form of *coarse graining*: the underlying points in the continuous phase space are the microstates, and the partitioning defines macrostates. Discrete systems can be similarly coarse-grained.

B.3 Coarse graining and emergence

[Ryan 2006] argues that emergent properties are simply a difference between global and local structure, rather than anything to do with 'levels'. He defines emergence in terms of *scope* (what is inside or outside the system boundary), *resolution* (fine or coarse-grained spatial and temporal distinctions; finer grained often implies narrower scope), and *state* (the information that distinguishes system configurations, at the given resolution). Fine grain and narrow scope are properties of *microstates*, coarser grains and wider scope of *macrostates*. He states that: *A property is emergent iff it is present in a macrostate and it is not present in the microstate*. He argues that this cannot be achieved by a change of resolution alone: it requires a change of scope. (Much of his argument compares scope and resolution, ignoring the state aspect.)

In our approach L can be considered to be finer-grained resolution microstates; S the coarser-grained macrostate. However, our main requirement is for a change of language between the levels (which may relate to [Ryan 2006]'s state information).

[Shalizi & Moore 2006] define "*a relation of 'emergence' between two sets of causal variables if (1) one is a coarse-graining of the other and (2) the coarse-grained variables can be predicted more efficiently*". Thermodynamics as a coarse graining of statistical mechanics is a classic example of this approach.

B.4 Example: coarse graining CAs

[Israeli & Goldenfeld 2006] give an example of coarse-graining cellular automata (CAs). The idea of the coarse graining is to capture the dynamics of a CA rule with another rule, expressed over a coarser grain in space and time, that preserves the underlying dynamics (although it will lose 'irrelevant' detail).

From one point of view, this might not seem to be an example of emergence: the two languages L and S are the same: that of CA rules. However, from another point of view, it can be considered to be emergence: the languages are different, in that they are *different* CA rules (except in a few cases) and they lose information (in that low-level fine-grained detail is washed out). Interestingly, the figures illustrating coarse grainings in [Israeli & Goldenfeld 2006] seem to highlight some of the underlying L structure (for

example, various propagating 'signals'), maybe because they have smoothed out other, irrelevant, structure.

The higher level rule S may be a lower complexity CA rule than the lower level L: because "*the system (the update rule, not the cell lattice) does not contain enough information to be complex at large scales*" [Israeli & Goldenfeld 2006]. This complexity-reducing kind of coarse graining (of 'relevant degrees of freedom') removes information that is relevant to the dynamics at L, but is redundant under the coarse graining at S: it does not capture all of the microscopic dynamics.

The coarse-graining provides an *exact* fit over all possible states: it loses information, but the systems never differ in their predictions. It is a form of information *compression*, not *approximation*.