Mathematics for Complex Systems:

The Objective Relativity of Complexity and Entropy

David Feldman

College of the Atlantic

and

The Santa Fe Institute

http://hornacek.coa.edu/dave/

Collaborator: Jim Crutchfield (UC Davis and SFI)

Thanks to: Carl McTague, Cosma Shalizi, Karl Young

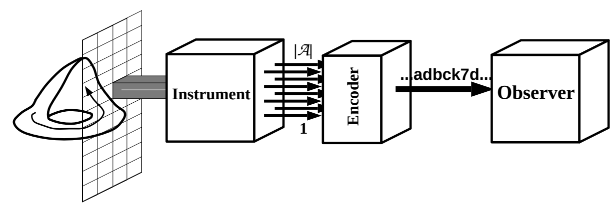
Overview and Motivation

- Complex systems pose a challenge for mathematics and mathematical sciences.
- Can mathematics be used at all for such systems? Or are such systems simply too complex to be simplified via mathematics?
- Central premise: the abstractions of mathematics and mathematical models can be used to gain qualitative insight into complex systems.
- In my remarks I will focus on two questions:
  1. What is complexity?
  2. What does it mean to model?
- I hope to convince you that the first question cannot be answered without answering the second question.

Why Complexity?

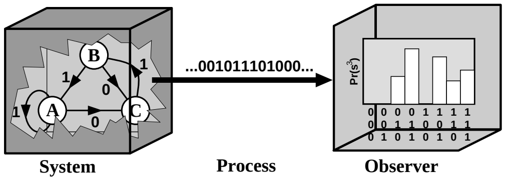
- Complexity is generally understood to be a measure of the difficulty of describing a thing or a process.
- There are many different contexts in which the term complexity is used:
  - Complexity as a measure of difficulty of learning a pattern (Bialek, et al, 2001)
  - Biological and ecological systems exhibit different levels of complexity and organization which we can study
  - Complexity(?) in evolution (McShea, 1991)
  - Complexity as measure of structure or pattern or correlation.
- I will focus on this last sort of complexity, but I think my general results extend to other types of complexity.

Measurement and Modeling



- On the left is “nature.”
- The act of measurement projects this system down to a lower dimension.
- These measurements are discretized.
- The measurements may then be encoded or corrupted by noise.
- They then reach the observer on the right, who wishes to make inferences about “nature.”
- Figure source: Crutchfield, 1992.

Modeling and Inference



- In very idealized form, the observer is faced with a long string of binary measurement data:

$$...0110101110101010111010110101011101110 1...$$

- What can the observer infer from this?
- The observer can determine the frequency (or probability) of occurrence of different sequences of 0's and 1's.
- Information theory gives us a way to measure properties of the sequence.

Shannon Entropy

Any time we use a probability distribution, this indicates some uncertainty. However, Some distributions indicate more uncertainty than others.

The Shannon Entropy H is the measure of the uncertainty associated with a probability distribution:

$$H[X] \equiv -\sum_x \Pr(x) \log_2 \Pr(x) \; . \qquad (1)$$

- A **Fair Coin**: (Probability of heads $= \frac{1}{2}$) has an unpredictability of 1.
- A **Biased Coin**: (Probability of heads $= 0.9$) has an unpredictability of 0.47.
- A **Perfectly Biased Coin**: (Probability of heads $= 1.0$) has an unpredictability of 0.00.

The conditional entropy is defined via:

$$H[X|Y] \equiv -\sum_x \Pr(x,y) \log_2 \Pr(x|y) \; . \qquad (2)$$

Entropy Rate

- The entropy rate h_μ is defined via

$$\lim_{L\to\infty} H[S_L | S_{L-1} S_{L-2} \ldots S_0] \; .$$

- In words: the entropy rate is the average uncertainty of the next symbol, given that an arbitrarily large number of symbols have already been seen.
- h_μ is the irreducible randomness: the randomness that persists even after statistics over arbitrarily long sequences are taken into account.
- h_μ is a measure of unpredictability.

Excess Entropy

- The excess entropy $\mathbf{E}$ is defined as the total amount of randomness that is "explained away" by considering larger blocks of variables.
- One can also show that $\mathbf{E}$ is equal to the mutual information between the "past" and the "future":

$$\mathbf{E} = I(\overrightarrow{S}; \overleftarrow{S}) \equiv H[\overrightarrow{S}] - H[\overrightarrow{S} \mid \overleftarrow{S}] \; .$$

- $\mathbf{E}$ is thus the amount one half "remembers" about the other, the reduction in uncertainty about the future given knowledge of the past.
- Equivalently, $\mathbf{E}$ is the "cost of amnesia:" how much more random the future appears if all historical information is suddenly lost.

Excess Entropy and Entropy Rate Summary

- Excess entropy $\mathbf{E}$ is a measure of complexity (order, pattern, regularity, correlation ...)
- Entropy rate h_μ is a measure of unpredictability.
- Both $\mathbf{E}$ and h_μ are well understood and have clear interpretations.
- For more, see, e.g., Grassberger 1986; Crutchfield and Feldman, 2003.
- I'll now consider 3 examples that illustrate some of the subtleties that are associated with measuring h_μ and $\mathbf{E}$.

Example I: Disorder as the Price of Ignorance

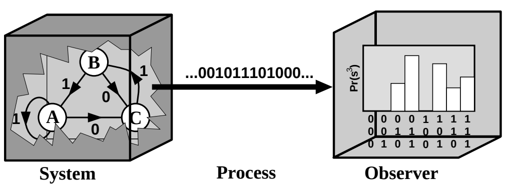
- Let us suppose that an observer seeks to estimate the entropy rate.
- To do so, it considers statistics over sequences of length L and then estimates h_μ using an estimator that assumes $\mathbf{E} = 0$.
- Call this estimated entropy ${h_\mu}'(L)$. Then, the difference between the estimate and the true h_μ is (Proposition 13, Crutchfield and Feldman, 2003):

$$h'_\mu(L) - h_\mu = \frac{\mathbf{E}}{L} \; . \tag{3}$$

- In words: The system appears more random than it really is by an amount that is directly proportional to the the complexity $\mathbf{E}$.
- In other words: regularities ($\mathbf{E}$) that are missed are converted into apparent randomness ($h'_\mu(L) - h_\mu$).

Example II: A Randomness Puzzle

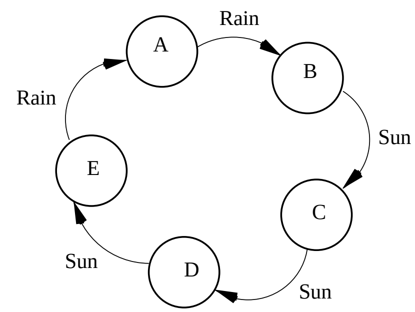
- Suppose we consider the binary expansion of π. Calculate its entropy rate h_μ and we'll find that it's 1.
- How can π be random? Isn't there a simple, deterministic algorithm to calculate digits of π?
- Yes. However, it is random if one uses histograms and builds up probabilities over sequences.
- This points out the *model-sensitivity* of both randomness and complexity.



- Histograms are a type of model. See, e.g., Knuth 2006.

Example III: Unpredictability due to Asynchrony

- Imagine a strange island where the weather repeats itself every 5 days. It's rainy for two days, then sunny for three days.



- You arrive on this deserted island, ready to begin your vacation. But, you don't know what day it is: $\{A, B, C, D, E\}$.
- Eventually, however, you will figure it out.

Example III: Unpredictability due to Asynchrony

- Once you are synchronized—you know what day it is—the process is perfectly predictable; $h_\mu = 0$.
- However, before you are synchronized, you are uncertain about the internal state. This uncertainty decreases, until reaching zero at synchronization.
- Denote by $\mathcal{H}(L)$ the average state uncertainty after L observations are made.
- The total state uncertainty experienced while synchronizing is the **Transient Information $\mathbf{T}$**:

$$\mathbf{T} \equiv \sum_{L=0}^{\infty} \mathcal{H}(L) \ . \qquad (4)$$

Example III: Unpredictability due to Asynchrony

- It turns out that different periodic sequences with the same P can have very different $\mathbf{T}$'s.
- For a given period P:

$$\mathbf{T}_{\max} \sim \frac{P}{2} \log_2 P \ , \qquad (5)$$

and

$$\mathbf{T}_{\min} \sim \frac{1}{2} \log_2^2 P \ , \qquad (6)$$

- E.g., if $P = 256$, then

$$\mathbf{T}_{\max} \approx 1024 \ , \ \text{and} \ \ \mathbf{T}_{\min} \approx 32 \ . \qquad (7)$$

- For much more, see Feldman and Crutchfield 2004.

Summary of Examples

- In all cases choice of representation and the state of knowledge of the observer influence the measurement of entropy or complexity.
  1. Ignored complexity is converted to entropy.
  2. π appears random.
  3. A periodic sequence is unpredictable.
- Hence, statements about unpredictability or complexity are necessarily a statement about the observer, the observed, and the relationship between the two.
- So complexity and entropy are relative, but in an objective, clearly specified way.

Conclusion: Modeling Modeling

- I have aimed to present an abstraction of the modeling process itself.
- These examples provide a crisp setting in which one can explore trade-offs between, say, the complexity of a model and the observed unpredictability of the object under study.
- The choice of model can strongly influence the result yielded by the model. This influence can be understood.
- The hope is these models of modeling can give us some general, qualitative insight into modeling.
- In my view, to study complex systems we often need to refine existing mathematical techniques and broaden our scope. However, we do not need a new kind of science.

References

- W. Bialek, et al., "Predictability, Complexity, and Learning." *Neural Computation*, 13:2409. 2001
- J.P. Crutchfield, "Knowledge and Meaning ... Chaos and Complexity." In Modeling Complex Systems. L. Lam and H.C. Morris, eds. Springer-Verlag. 66-10. 1992.
- J.P. Crutchfield, "The Calculi of Emergence." *Physica D*. 75:11. 1994.
- J.P. Crutchfield and D.P. Feldman, "Regularities Unseen, Randomness Observed." *Chaos*. 15:23-54. 2003.
- D.P. Feldman and J.P. Crutchfield, "Measures of Statistical Complexity, Why?" *Phys. Lett. A*, 238:244. 1998.
- D.P. Feldman and J.P. Crutchfield, "Synchronizing to Periodicity." *Advances in Complex Systems.* 7:329-355. 2004.

- P. Grassberger. "Toward a Quantitative Theory of Self-Generated Complexity." *Intl. J. Theo. Phys.* 25:907. 1986.
- K.H. Knuth. "Optimal Data-Based Binning for Histograms." Pre-print. http://arxiv.org/abs/physics/0605197. 2006.
- D.W. McShea. "Complexity and evolution: what everybody knows." Biology and Philosophy 6:303-324. 1991.
- C.R. Shalizi. "Methods and Techniques of Complex Systems Science: An Overview." in *Complex Systems Science in Biomedicine*, T.E. Deisboeck, et al (eds.), Kluwer. 2004.