An Incomplete Survey of Complexity Measures and Some Additional Thoughts on Complexity

The term *complexity* has many different meanings. At least one adjective is needed to help distinguish between different uses of the word:

- Kolmogorov-Chaitin Complexity
- Computational Complexity
- Stochastic Complexity
- Statistical Complexity
- Structural Complexity
-
-

Note: Some portions of this presentation were prepared jointly with Jim Crutchfield.

Deterministic Complexity

The *Kolmogorov-Chaitin* complexity $K(x)$ of an object x is the length, in bits, of the smallest program (in bits) that when run on a *Universal Turing Machine* outputs x and then halts.

References:

- Kolmogorov, *Problems of Information Transmission*, 1:4-7. (1965)
- Kolmogorov, *IEEE Trans. Inform. Theory*, IT-14:662-664. (1968).
- Solomonoff. *Inform. Contr.*, 7:1-22, 224-254. (1964).
- Chaitin, *J. Assoc. Comp. Mach.*, 13:547-569. (1966).
- Martin-Löf, *Inform. Contr.*, 9:602-619. (1966).
- **Books:**
  - Chpt. 7 of: Cover and Thomas, “Elements of Information Theory,” Wiley, 1991.
  - Chaitin, “Information, Randomness and Incompleteness,” World Scientific, 1987.

Kolmogorov Complexity $\approx$ Randomness

- The Kolmogorov complexity $K(x)$ is maximized for random strings, since it requires a deterministic accounting of all symbols in the string.
- The average growth rate of $K(x)$ is equal to the entropy rate h_μ.
- If x = trajectory of a chaotic dynamical system f:

$$K(x(t)) \;=\; h_\mu(f) \quad \text{for typical } x(0) \; .$$

(Brudno, *Trans. Moscow Math. Soc.*, 44:127. (1983).)

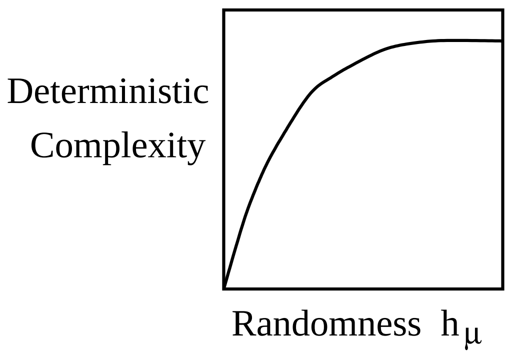
- If a string x is random, then it possesses no regularities. Thus,

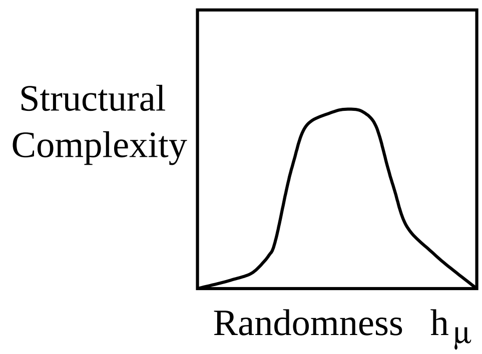
$$K(x) \;=\; |\mathrm{Print}(x)| \; .$$

- That is, the shortest program to get a UTM to produce x is to just hand the computer a copy of x and say “print this.”

Measures of "Complexity" that Capture a Property Distinct from Randomness

- The entropy rate h_μ and the Kolmogorov Complexity $K(x)$ do not measure pattern or structure or correlation or organization.
- h_μ and $K(x)$ are maximized for random strings.





- Structural complexity or statistical complexity measures are not maximized by random strings.
- As we've seen, there are many possible behaviors for structural complexity as a function of entropy.
- The excess entropy and the statistical complexity are examples of structural complexity measures.
- The following slides review a few other statistical complexity measures.

Other Approaches to Structural Complexity

Logical Depth:

The **Logical Depth** of x is the **run time** of the shortest program that will cause a UTM to produce x and then halt.

Logical depth is not a measure of randomness; it is small for both trivially ordered and random strings.

References:

- Bennett, *Found. Phys.*, 16:585-592, 1986.
- Bennett, in *Complexity, Entropy and the Physics of Information*, Addison-Wesley, 1990.

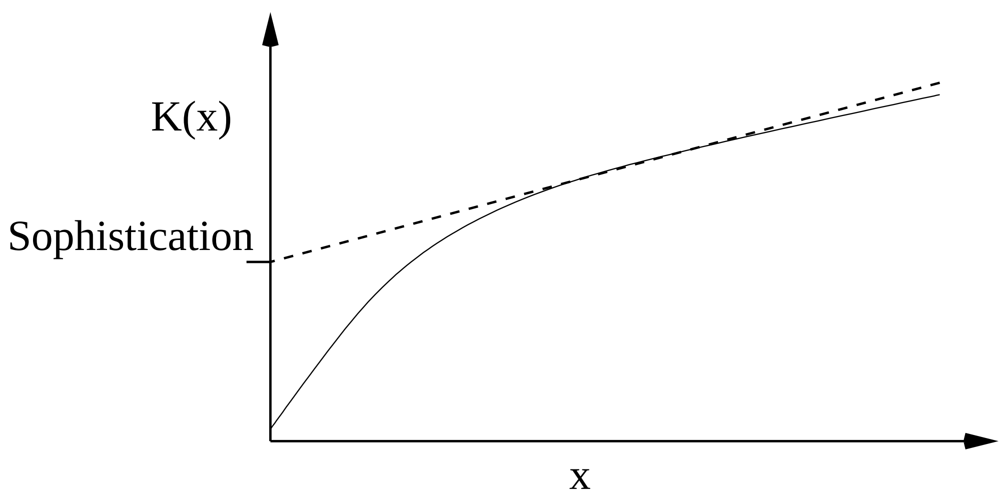
Thermodynamic Depth:

Proposed as a measure of structural complexity (Lloyd and Pagels, *Annals of Physics*, 188:186-213). Thermodynamic depth of an object is the total entropy generated in the production of that object. However, thermo. depth depends crucially on the choice of state. Lloyd and Pagels give no general prescription for how states should be chosen. Once states are chosen, thermo. depth is equivalent to the reverse time entropy rate. (Shalizi and Crutchfield, Physical Review E 59 1999.)

Other Approaches to Structural Complexity, continued

Sophistication:

- Koppel, *Complex Systems*, 1:1087-91, 1987.
- The Kolmogorov complexity $K(x)$ of an object will grow linearly in x. The sophistication is essentially the size of the model—the part of $K(x)$ that doesn't grow linearly:



- The linear growth is that portion of $K(x)$ that describes the length of the string.
- The constant part is the portion of $K(x)$ that describes the regularities of the string.
- Roughly speaking, the sophistication can be thought of as the Kolmogorov version of the excess entropy.

Other Approaches to Structural Complexity, continued

Non-Linear Modeling

- Wallace and Boulton, 1968.
- Crutchfield and McNamara, *Complex Systems* 1: 417-452, 1987.
- Rissanen, *Stochastic Complexity in Statistical Inquiry*, World Scientific, 1989.
- Crutchfield and Young, in *Complexity, Entropy and the Physics of Information*, Addison-Wesley, 1990.

Model Convergence and Hierarchical Grammatical Complexities:

- Badii and Politi, *Complexity: Hierarchical Structures and Scaling in Physics*, Cambridge, 1997.
- Badii and Politi, *Phys. Rev. Lett.*, 78:444-447, 1997.

Note: Badii and Politi's book contains a solid discussion of many different structural complexity measures.

Other Approaches to Structural Complexity, continued

Miscellaneous References:

- Kolmogorov, *Russ. Math. Surveys*, 38:29, 1983.
- Wolfram, *Comm. Math. Phys.*, 96:15-57, 1984.
- Wolfram, *Physica D*, 10:1-35, 1984.
- Hubermann and Hogg, *Physica D*, 22:376-384, 1986.
- Bachas and Hubermann, *Phys. Rev. Lett.*, 57:1965, 1986
- Peliti and Vulpiani, eds., *Measures of Complexity*, Springer-Verlag, 1988.
- Wackerbauer, et. al., *Chaos, Solitons & Fractals*, 4:133-173, 1994.
- Bar-Yam, *Dynamics of Complex Systems*, Addison-Wesley, 1997.

Non-constructive Complexity: The Road Untakable

All Universal Turing Based complexity measures suffer from several drawbacks:

1. They are uncomputable.
2. By adopting a UTM, the most powerful discrete computation model, one loses the ability to distinguish between systems that can be described by computational models less powerful than a UTM.

UTM-based "complexity" measures include:

- **Logical Depth:** Bennett, *Found. Phys.*, 16:585-592, 1986.
- **Sophistication:** Koppel, *Complex Systems*, 1:1087-91, 1987.
- **Effective Complexity:** Gell-Mann and Lloyd, *Complexity*, 2:44-52, 1996.

On the other hand UTM-based arguments are useful for providing a clear framework for expressing notions of complexity.

Complexity = Order $\times$ Disorder?

- There are a number of complexity measures of the form:

$$\text{Complexity} = \text{Order} \times \text{Disorder}$$

- Disorder is usually some form of entropy.
- Sometimes "order" is simply $(1 - h_\mu)$.
- Often, "order" is taken to be some measure of "distance from equilibrium," where equilibrium and equiprobability are sometimes considered to be synonymous.

In my view these sorts of complexity measures have some serious shortcomings:

- Lack a clear interpretation and direct accounting of structure.
- Unclear that distance from equilibrium is equivalent to order.
- Assign a value of zero complexity to all systems with vanishing entropy.

Complexity = Order $\times$ Disorder?, continued

But, you can read the papers and decide for yourself. See, e.g.,

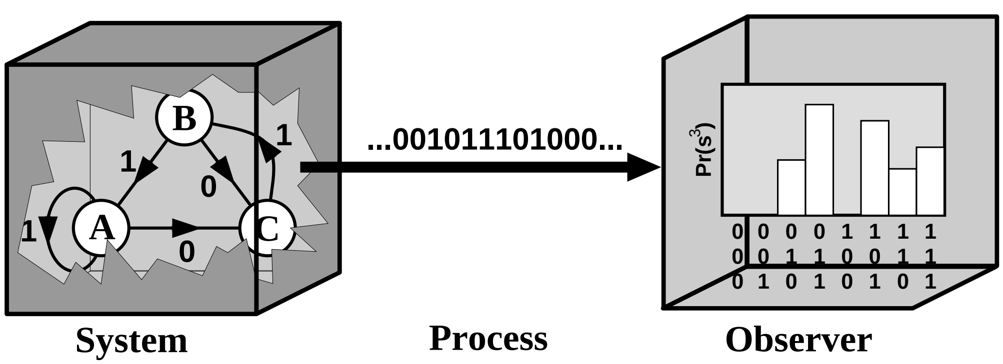
- Shiner, et al. *Phys. Rev. E*, 59:1459. 1999.
- Lopez-Ruiz, et al., *Phys. Lett. A*, 209:321. 1995.
- Piasecki, et al., *Physica A*, 307:157. 2002.

For some critiques, see:

- Feldman and Crutchfield, *Phys. Lett. A*, 238:244. 1998.
- Crutchfield, et al., *Phys. Rev. E*, 62:2996. 2000.
- Binder. *Phys. Rev. E*, 62:2998. 2000.

A randomness puzzle

- Suppose we consider the binary expansion of π. Calculate it's entropy rate h_μ and we'll fi nd that it's 1.
- How can π be random?
- It is not random if one uses Kolmogorov complexity, since presumably there's a fairly short algorithm to produce the digits of π.
- But 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.

Model Dependence

- There is no (computable), all-purpose measure of randomness or complexity.
- This isn't cause for despair. Just be as clear as you can about your modeling assumptions.
- Sometimes modeling assumptions can be hidden.
- I don't think will ever be a 100% objective measure of complexity. A statement about complexity will always be, to some extent, a statement about both the observer and the observed.

Early Uses of Mutual Information

- Rothstein, in *The Maximum Entropy Formalism*, MIT Press, 1979.
- Chaitin, in *Information, Randomness, and Incompleteness*, World Scientifi c, 1987.
- Gatlin, *Information Theory and the Living System*, Columbia University Press. 1972.
- Watanabe, *Knowing and Guessing: A Quantitative Study of Inference and Information*, Wiley, 1969.