A Survey of
“Complexity Measures”

11 June 1998

David P. Feldman
University of California, Davis
dave@santafe.edu
http://leopard.ucdavis.edu/dave/

and

Jim Crutchfield
The Santa Fe Institute
chaos@santafe.edu
http://santafe.edu/˜jpc/

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
-
-

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.

Measures of Randomness

The **entropy rate** h_μ of a symbolic sequence measures the unpredictability (in bits per symbol) of the sequence.
The entropy rate is also known as the **entropy density** or the **metric entropy**.

References:

- Boltzmann (1866).
- Shannon, *Bell Sys. Tech. J.* 27:379-423. (1948).
- Kolmogorov, *Dolk. Akad. Nauk. SSSR*, 119:861-864. (1958).
- **Books:**
  - Shannon and Weaver, "The Mathematical Theory of Communication," Univ. of Illinois Press, (1963).
  - Rényi, "Probability Theory," North Holland, 1970.
  - Chpt. 7 of: Cover and Thomas, "Elements of Information Theory," Wiley, 1991.
  - Beck and Schlögl, "Thermodynamics of Chaotic Systems," Cambridge, (1993).

Kolmogorov Complexity ≈ Randomness!!

The Kolmogorov complexity $K(x)$ is maximized for random strings.

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).)

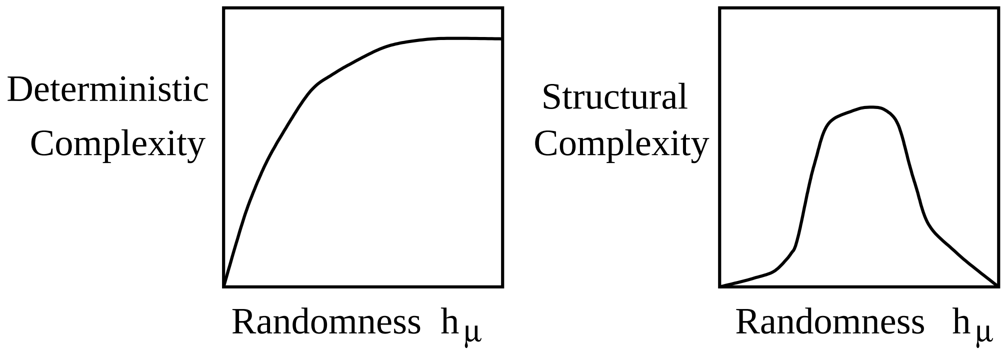
If a string x is random, then it possesses no regularities. Thus,

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

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.



Structure or pattern is maximized for neither high nor low randomness.

Note: The structural complexity vs. randomness relation above is just one of many possible behaviors. Different systems have different structural complexity vs. randomness plots. **There is no "universal" complexity-entropy relationship!** (E.g., Feldman and Crutchfield, *Phys. Lett. A*, 238:244-252, (1998), and references therein.)

Information Theoretic Approaches to Structural Complexity

Entropy Density Convergence and/or Mutual Information:

- Crutchfield and Packard, *Intl. J. Theo. Phys*, 21:433-466. (1982); *Physica D*, 7:201-223, 1983.
- Shaw, “The Dripping Faucet ..., ” Aerial Press, 1984.
- Grassberger, *Intl. J. Theo. Phys*, 25:907-938, 1986.
- Szépfalusy and Györgyi, *Phys. Rev. A*, 33:2852-2855, 1986.
- Lindgren and Nordahl, *Complex Systems*, 2:409-440. (1988).
- Csordás and Szépfalusy, *Phys. Rev. A*, 39:4767-4777. 1989.
- Li, *Complex Systems*, 5:381-399, 1991.
- Freund, Ebeling, and Rateitschak, *Phys. Rev. E*, 54:5561-5566, 1996.
- Feldman and Crutchfield, *J. Stat. Phys* (submitted) SFI:98-04-026, 1998.

Entropy Density Convergence and/or Mutual Information
Notes on Terminology

All of the following terms refer to (essentially) the same quantity.

- **Excess Entropy:** Crutchfield, Packard, Feldman
- **Stored Information:** Shaw
- **Effective Measure Complexity:** Grassberger, Lindgren, Nordahl
- **Reduced (Rényi) Information:** Szépfalusy, Györgyi, Csordás
- **Complexity:** Li, Arnold

Early Uses of Mutual Information

- Rothstein, in *The Maximum Entropy Formalism*, MIT Press, 1979.
- Chaitin, in *Information, Randomness, and Incompleteness*, World Scientific, 1987.
- Watanabe, *Knowing and Guessing: A Quantitative Study of Inference and Information*, Wiley, 1969.

Computational Mechanics

Discover and Quantify Structure by Using a Combination of Computation Theory, Statistical Inference, and Information Theory

Computational Mechanics seeks to Detect the Intrinsic Computation being Performed by the System

- Crutchfield and Young, *Phys. Rev. Lett*, 63:105-108, 1989
- Crutchfield and Young, in *Complexity, Entropy and the Physics of Information*, Addison-Wesley, 1990.
- Crutchfield, *Physica D*, 75:11-54, 1994.
- Hanson, *PhD Thesis*, University of California, Berkeley, 1993.
- Hanson and Crutchfield, *Physica D*, 103:169-189, 1997.
- Upper, *PhD Thesis*, University of California, Berkeley, 1997.
- Delgado and Solé, *Phys. Rev. E*, 55:2338-2344, 1997.
- Witt, Neiman and Kurths, *Phys. Rev. E*, 55:5050-5059, 1997.
- Goncavales, et. al., *Physica A*, (in press), 1998.
- Feldman and Crutchfield, *J. Stat. Phys* (submitted) SFI:98-04-026, 1998.

Other Approaches to Complexity Distinct from Randomness

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). 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, (in preparation), 1998.)

Other Approaches to Complexity Distinct from Randomness

Sophistication:

Reference:

- Koppel, *Complex Systems*, 1:1087-91, 1987.

Effective Complexity:

Reference:

- Gell-Mann and Lloyd, *Complexity*, 2:44-52, 1996.

Other Approaches to Complexity Distinct from Randomness

Non-Linear Modeling

References:

- 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:

References:

- 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 Complexity Distinct from Randomness

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 Measures: 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.

Falling off the "Edge of Chaos"

- Packard, "Adaptation to the Edge of Chaos" in *Dynamic Patterns in Complex Systems*, Kelso et.al, eds., World Scientific, 1988
- Mitchell, Hraber, and Crutchfield "Revisiting the 'Edge of Chaos' *Complex Systems*, 7:89-130, 1993. (Rebuttal to Packard, 1988).
- Crutchfield and Young, "Inferring Statistical Complexity", *Phys. Rev. Lett.*, 63:105-108, 1989. (First plot of "Complexity" vs. Entropy.)
- Langton "Computation at the Edge of Chaos", *Physica* D (1990).
- Li, Packard and Langton, "Transition Phenomena in Cellular Automata Rule Space" *Physica D* 45 (1990) 77.
- Wooters and Langton, "Is there a Sharp Phase Transition for Deterministic Cellular Automata?", *Physica D* 45 (1990) 95.
- Crutchfield, "Unreconstructible at Any Radius", *Phys. Lett.* A 171: 52-60, 1992.

This paper surveys and analyzes measures which can be extracted from the training datasets in order to characterize the complexity of the respective classification problems. Their use in recent literature is also reviewed and discussed, allowing to prospect opportunities for future work in the area. Finally, descriptions are given on an R package named Extended Complexity Library (ECoL) that implements a set of complexity measures and is made publicly available. Discover the world's research. 15+ million members. Particularly, code complexity measures have been designed to measure and monitor complexity in practice (Zuse 1991; Abran 2010; Fenton and Bieman 2014). Complexity measurement allows quantifying complexity and understanding its effect on maintainability and defect proneness. The concept of code complexity, however, is not an atomic concept, so it is difficult to design a single measure that quantifies code complexity thoroughly. Instead, several complementary measures are designed to measure different aspects of complexity.Â At the end of this part of the survey, an open question was included to allow respondents to suggest other code characteristics that they believed could significantly increase complexity. 3.3 Complexity and Internal Code Quality Attributes. indirect measure of well complexity, since complex wells are frequently associated. with multiple strings, and narrow margins between formation pore pressure and. fracture pressure gradients often result in the requirement of a greater number of.Â The key drilling factors are user-defined qualitative variables %i that are assigned. an integer-valued weight Ïˆ i (w) according to the occurrence of the condition and its. degree of complexity. Let Ïˆ i denote the ith drilling factor of well w and Ïˆ i (w) the. corresponding numerical weight