

# An introduction to information theory and entropy

Tom Carter

<http://cogs.csustan.edu/~tom/SFI-CSSS>

Complex Systems Summer School

June, 2002

# Our general topics: ←

- ⊙ Measuring complexity
- ⊙ Some probability background
- ⊙ Basics of information theory
- ⊙ Some entropy theory
- ⊙ The Gibbs inequality
- ⊙ A simple physical example (gases)
- ⊙ Shannon's communication theory
- ⊙ Application to Biology (analyzing genomes)
- ⊙ Some other measures
- ⊙ Some additional material
- ⊙ Examples using Bayes' Theorem
- ⊙ Analog channels
- ⊙ Application to Physics (lasers)
- ⊙ References

# The quotes



- ⊙ Science, wisdom, and counting
- ⊙ Being different – or random
- ⊙ Surprise, information, and miracles
- ⊙ Information (and hope)
- ⊙ H (or S) for Entropy
- ⊙ Thermodynamics
- ⊙ Language, and putting things together
- ⊙ Tools

# Science, wisdom, and counting ←

“Science is organized knowledge. Wisdom is organized life.”

- Immanuel Kant

“My own suspicion is that the universe is not only stranger than we suppose, but stranger than we can suppose.”

- John Haldane

“Not everything that can be counted counts, and not everything that counts can be counted.”

- Albert Einstein (1879-1955)

“The laws of probability, so true in general, so fallacious in particular .”

- Edward Gibbon

# Measuring complexity ←

- Workers in the field of complexity face a classic problem: how can we tell that the system we are looking at is actually a complex system? (i.e., should we even be studying this system? :-)

Of course, in practice, we will study the systems that interest us, for whatever reasons, so the problem identified above tends not to be a real problem. On the other hand, having chosen a system to study, we might well ask “How complex is this system?”

In this more general context, we probably want at least to be able to compare two systems, and be able to say that system A is more complex than system B.

Eventually, we probably would like to have some sort of numerical rating scale.

- Various approaches to this task have been proposed, among them:
  1. Human observation and (subjective) rating
  2. Number of parts or distinct elements (what counts as a distinct part?)
  3. Dimension (measured how?)
  4. Number of parameters controlling the system
  5. Minimal description (in which language?)
  6. Information content (how do we define/measure information?)
  7. Minimal generator/constructor (what machines/methods can we use?)
  8. Minimum energy/time to construct (how would evolution count?)

- Most (if not all) of these measures will actually be measures associated with a *model* of a phenomenon. Two observers (of the same phenomenon?) may develop or use very different models, and thus disagree in their assessments of the complexity. For example, in a very simple case, counting the number of parts is likely to depend on the scale at which the phenomenon is viewed (counting atoms is different from counting molecules, cells, organs, etc.).

We shouldn't expect to be able to come up with a single universal measure of complexity. The best we are likely to have is a measuring system useful by a particular observer, in a particular context, for a particular purpose.

My first focus will be on measures related to how surprising or unexpected an observation or event is. This approach has been described as *information theory*.

# Being different – or random



“The man who follows the crowd will usually get no further than the crowd. The man who walks alone is likely to find himself in places no one has ever been before. Creativity in living is not without its attendant difficulties, for peculiarity breeds contempt. And the unfortunate thing about being ahead of your time is that when people finally realize you were right, they’ll say it was obvious all along. You have two choices in life: You can dissolve into the mainstream, or you can be distinct. To be distinct is to be different. To be different, you must strive to be what no one else but you can be. ”

-Alan Ashley-Pitt

“Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin.”

- John von Neumann (1903-1957)