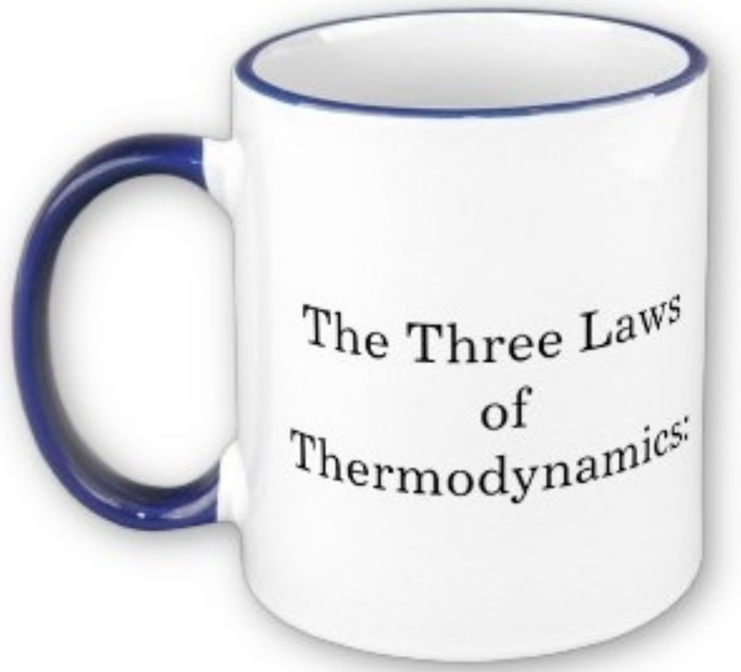
Information a condition of self-organization

How information is represented, communicated and processed in complex systems…

> "Although [complex systems] differ widely in their physical attributes, they resemble one another in the way they handle information. That common feature is perhaps the best starting point for exploring how they operate."
>
> —— Murray Gell-Mann, *The Quark and the Jaguar*, 1995

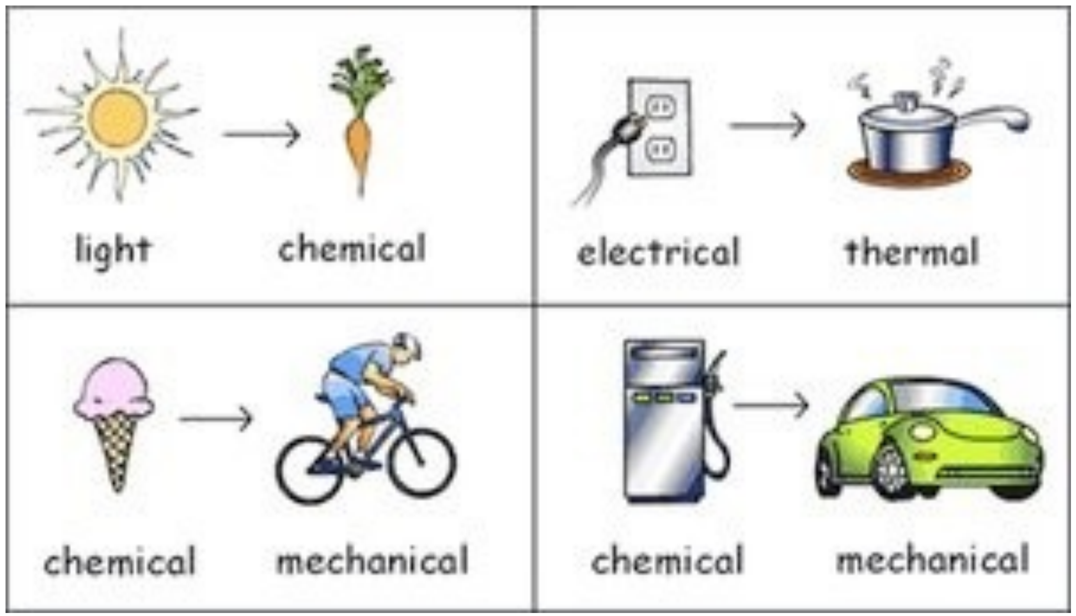
The mathematical structure information starts with...



First law of thermodynamics: In an isolated system, energy is conserved

Energy: A system's potential to do "work"

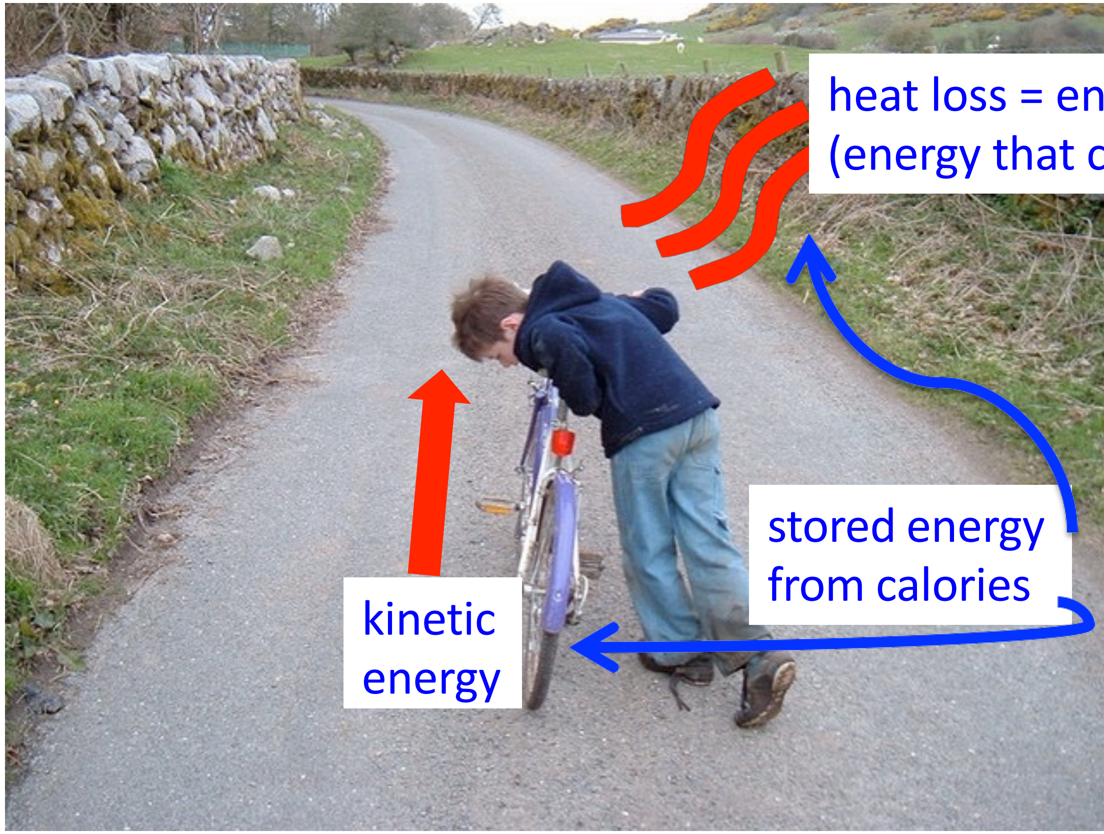
Energy can be transformed from one kind into another:



http://www.eoearth.org/article/AP_Environmental_Science_Chapter_1-_Flow_of_Energy

The total amount of energy in the system always remains the same.

Second law of thermodynamics: In an isolated system, entropy always increases until it reaches a maximum value.



Energy that doesn't contribute to moving the bike, this is entropy produced by the transfer energy.

Entropy can be thought of as a measure of disorder of a system

http://www.flickr.com/photos/ztephen/130961009/sizes/z/in/photostream/

Implications of the Second Law of Thermodynamics

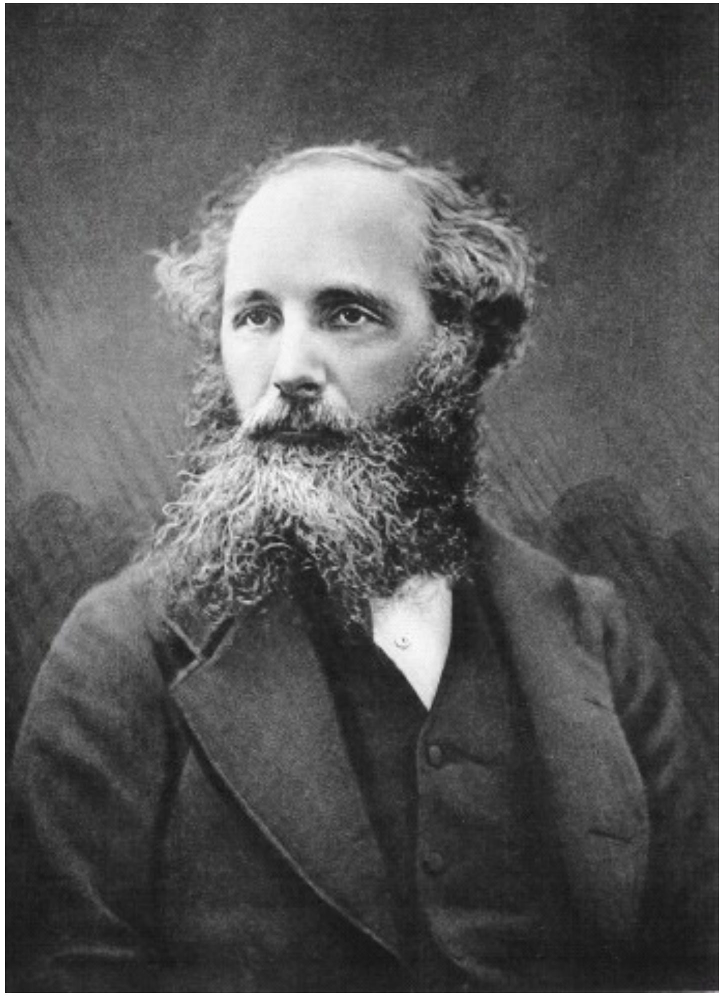
- Systems are naturally disordered. They cannot become organized without the input of work

- Perpetual motion machines are not possible

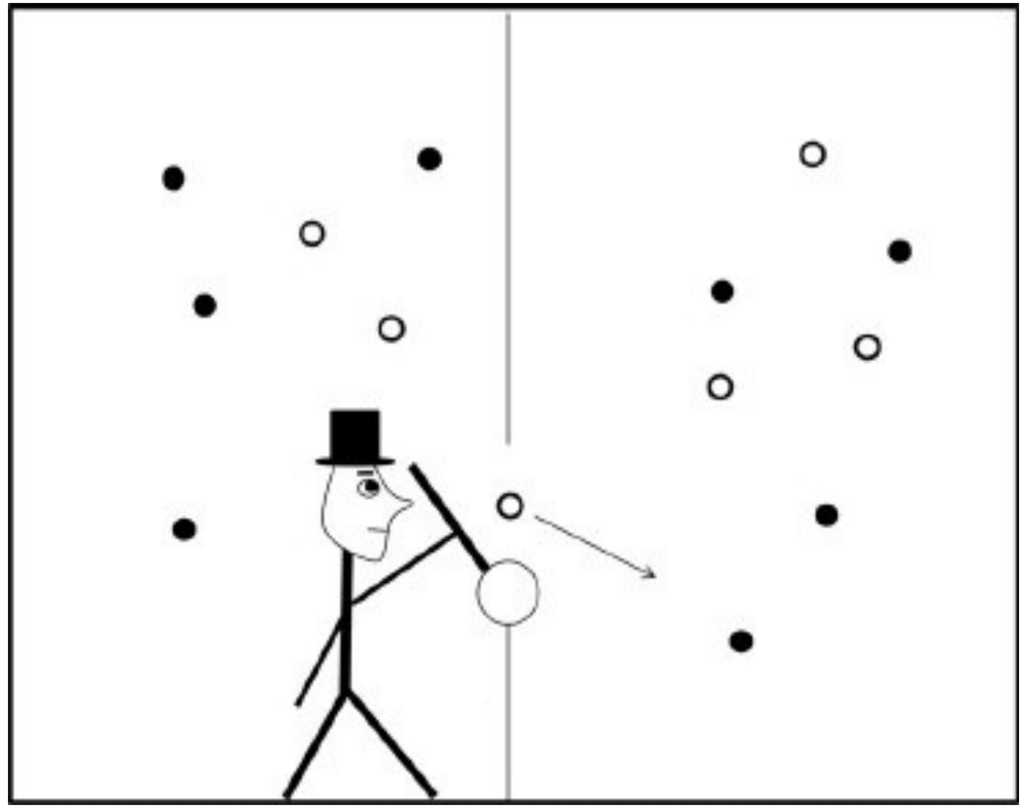
- Time has a direction: the direction of increasing entropy

Disorder in a system always increase until it reaches its maximum value

Paradox of Maxwell´s Demon



James Clerk Maxwell, 1831-1879



Maxwell’s Demon

“The hot system [i.e., the right side] has gotten hotter and the cold [the left side] has gotten colder and yet no work has been done, only the intelligence of a very observant and neat-fingered being has been employed.”

Maxwell:

The second law of thermodynamics is “a statistical certainty”.

THEORY OF HEAT

BY

J. CLERK MAXWELL, M.A.
LL.D. Edin., F.R.SS. L. & E.
Honorary Fellow of Trinity College
Professor of Experimental Physics in the University of Cambridge

WITH CORRECTIONS AND ADDITIONS (1891)
BY

LORD RAYLEIGH, M.A., D.C.L., LL.D.
Secretary of the Royal Society, Professor of Natural Philosophy in the Royal Institution, and late Professor of Experimental Physics in the University of Cambridge

NEW IMPRESSION

LONGMANS, GREEN, AND CO.
39 PATERNOSTER ROW, LONDON
NEW YORK AND BOMBAY
1902

Leo Szilard, 1898-1964

ON THE DECREASE OF ENTROPY IN A THERMODYNAMIC SYSTEM BY THE INTERVENTION OF INTELLIGENT BEINGS

LEO SZILARD

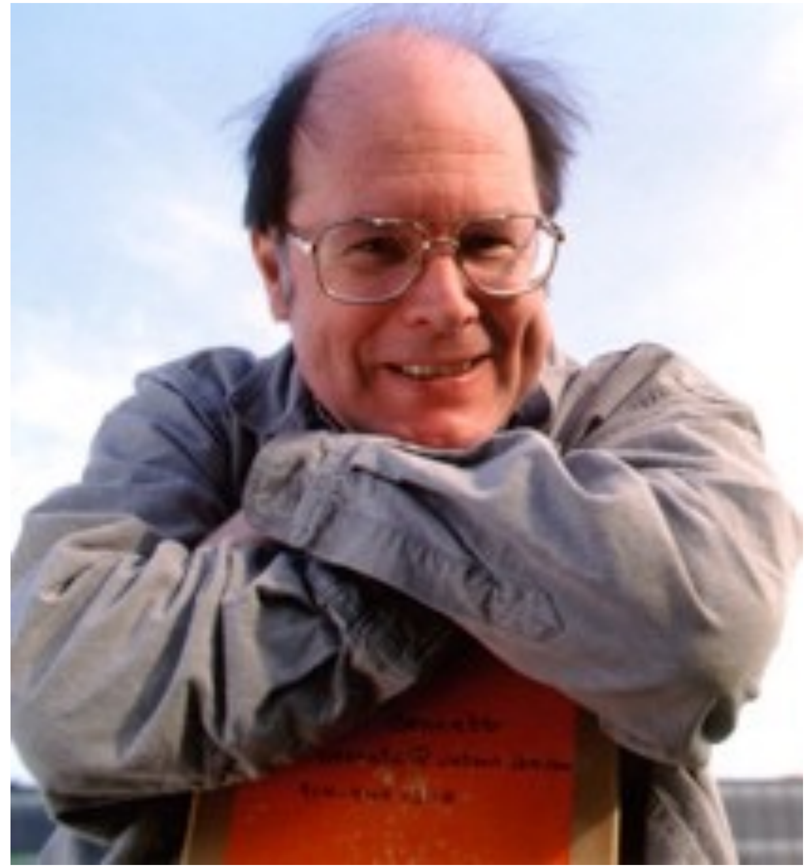
Translated by Anatol Rapoport and Mechthilde Knoller from the original article "Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen." Zeitschrift für Physik, 1929, 53, 840–856.

A “bit” of information

Szilard: A *bit* of information is the amount of information needed to answer a “fast/slow” question, or any “yes/no” question.

The field of Computer Science adopted this terminology for computer memory.

Rolf Landauer (1927−1999)



Charles Bennett

The Physics of information… the
information is a physical property…

ELSEVIER

15 July 1996

PHYSICS LETTERS A

Physics Letters A 217 (1996) 188–193

The physical nature of information

Rolf Landauer [1]

IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, USA

Received 9 May 1996
Communicated by V.M. Agranovich

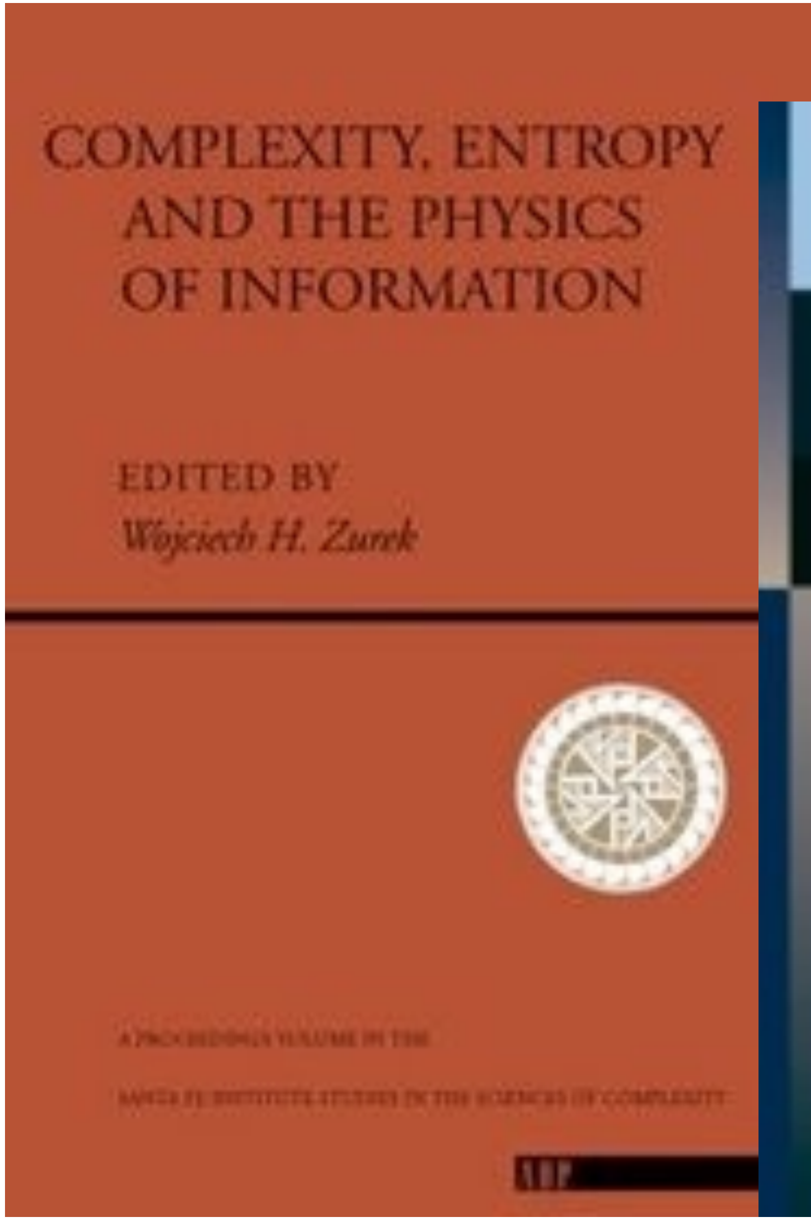
Abstract

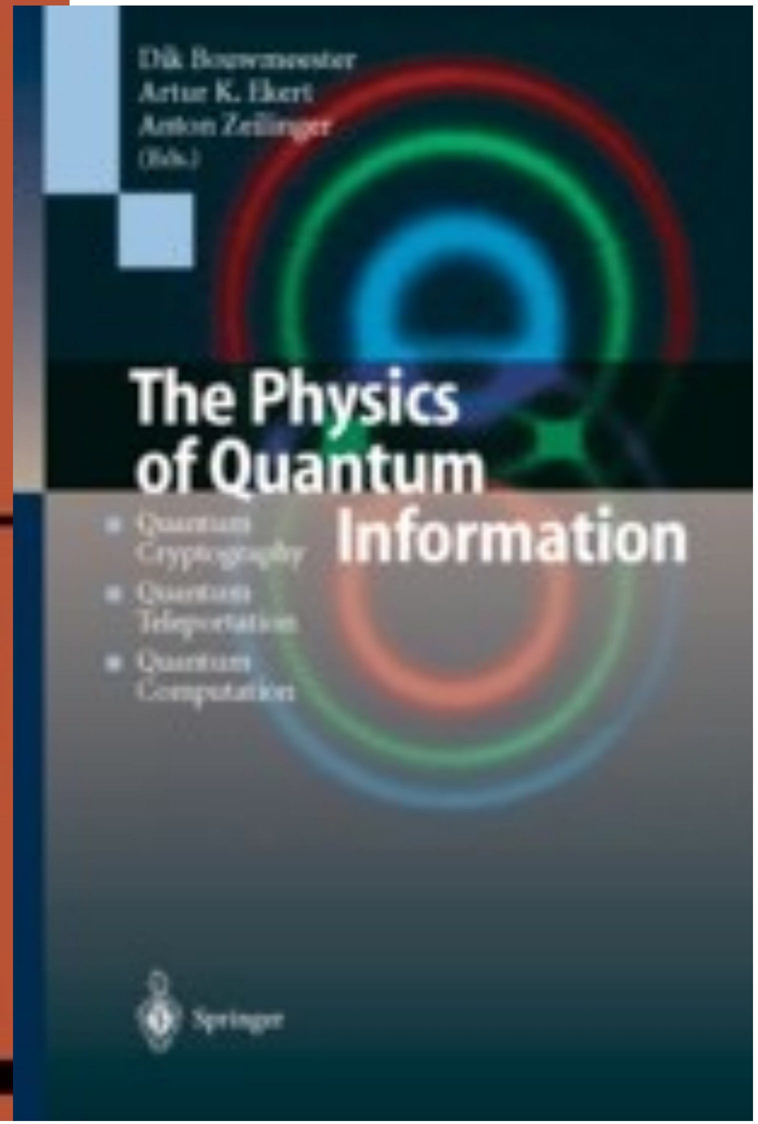
Information is inevitably tied to a physical representation and therefore to restrictions and possibilities related to the laws of physics and the parts available in the universe. Quantum mechanical superpositions of information bearing states can be used, and the real utility of that needs to be understood. Quantum parallelism in computation is one possibility and will be

The Fundamental Physical Limits of Computation

What constraints govern the physical process of computing? Is a minimum amount of energy required, for example, per logic step? There seems to be no minimum, but some other questions are open

by Charles H. Bennett and Rolf Landauer


COMPLEXITY, ENTROPY
AND THE PHYSICS
OF INFORMATION
EDITED BY
Wojciech H. Zurek


Dik Bouwmeester
Artur K. Ekert
Anton Zeilinger
(Eds.)
The Physics of Quantum Information
Quantum Cryptography
Quantum Teleportation
Quantum Computation
Springer

In electronics...

PHYSICAL REVIEW B **84**, 085418 (2011)

Probing the power of an electronic Maxwell's demon: Single-electron transistor monitored by a quantum point contact

Gernot Schaller,* Clive Emary, Gerold Kiesslich, and Tobias Brandes
Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstrasse 36, D-10623 Berlin, Germany
(Received 28 June 2011; revised manuscript received 14 July 2011; published 23 August 2011)

We suggest that a single-electron transistor continuously monitored by a quantum point contact may function as Maxwell's demon when closed-loop feedback operations are applied as time-dependent modifications of the tunneling rates across its junctions. The device may induce a current across the single-electron transistor even

In biology...

Life's demons: information and order in biology

What subcellular machines gather and process the information necessary to sustain life?

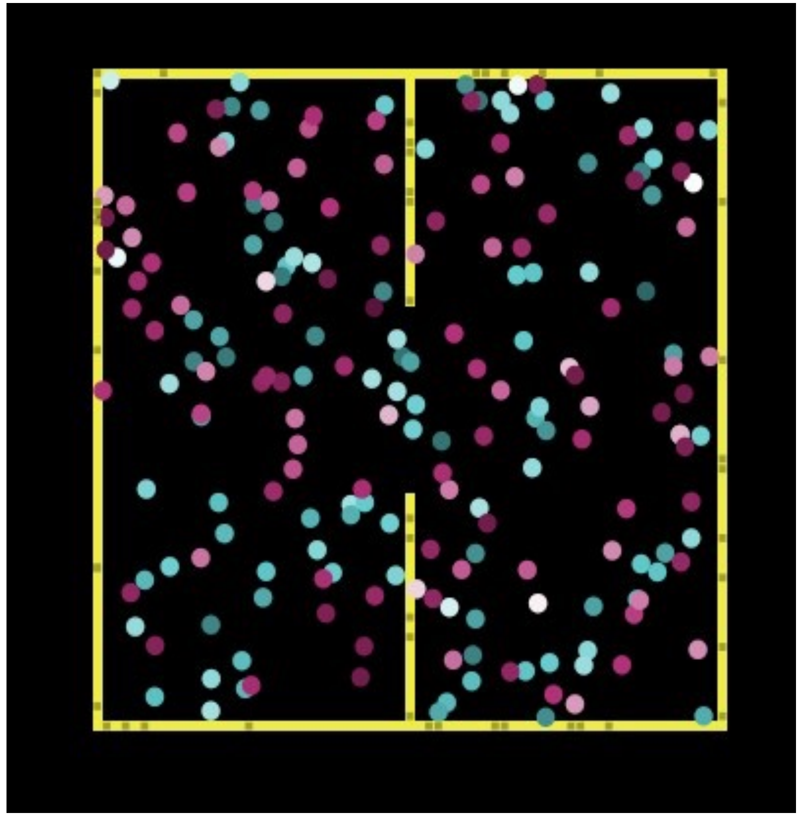
Philippe M Binder & Antoine Danchin

Received 8 February 2011; Accepted 15 April 2011

Thermodynamics: The study of heat and thermal energy

Statistical mechanics: A general mathematical framework that shows how macroscopic properties (e.g. heat) arise from statistics of the *mechanics* of large numbers of microscopic components (e.g., atoms or molecules)

Example: Room full of air



Macroscopic property (thermodynamics): Temperature, pressure

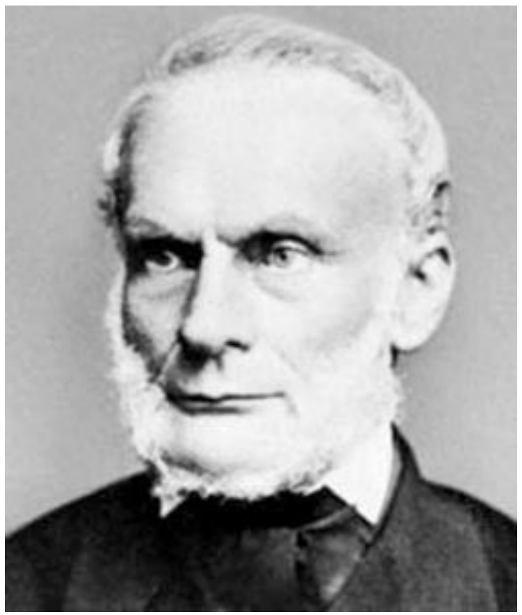
Microscopic property (mechanics): Positions and velocities of air molecules

Statistical mechanics: How statistics of positions and velocities of molecules give rise to temperature, pressure, etc.

Thermodynamic entropy measures the amount of heat loss when energy is transformed to work

Heat loss ≈ "disorder"

Theory is specific to heat



Rudolf Clausius, 1822-1888

Statistical mechanics entropy measures the number of possible microstates that lead to a macrostate

Number of microstates ≈ disorder

A more general theory

Ludwig Boltzmann, 1844-1906

A slight sidetrack to learn about microstates and macrostates

Microstates: specific state of the three slot-machine windows

Example microstate: {*cherry*, *lemon*, *apple*}

Note that a microstate here is a triple of fruit values, not a single fruit value. It is a description of the "state" of the slot machine.

Macrostate: Collection (or *set*) of microstates.

Example macrostate: *Win* (collection of microstates that have three of the same fruit showing).

M F P B L   M P F B L   H P F B L

5^3

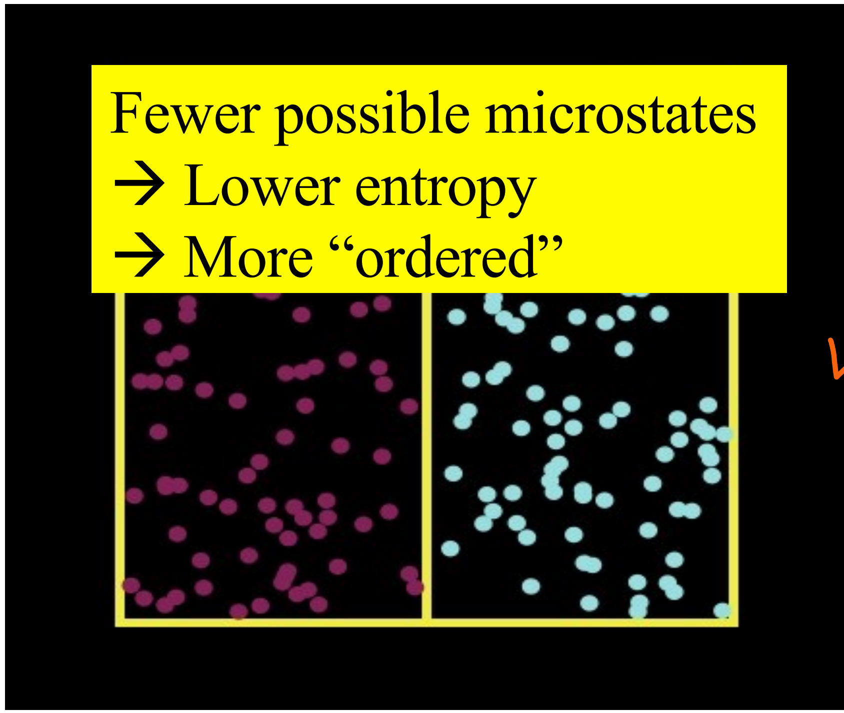
Question 1: How many microstates give rise to the *Win* macrostate? 5

Question 2: How many microstates give rise to the *Lose* macrostate? 120

NetLogo Two Gas Model

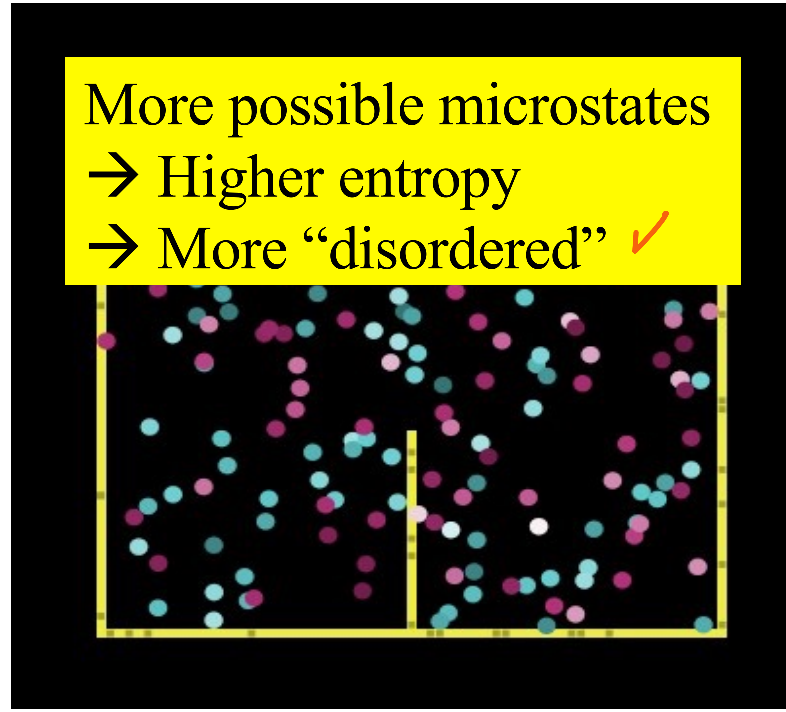
Microstate: Position and velocity of
of every particle

Start



Macrostate: All fast particles are on the right, all slow particles are on the left.

Finish

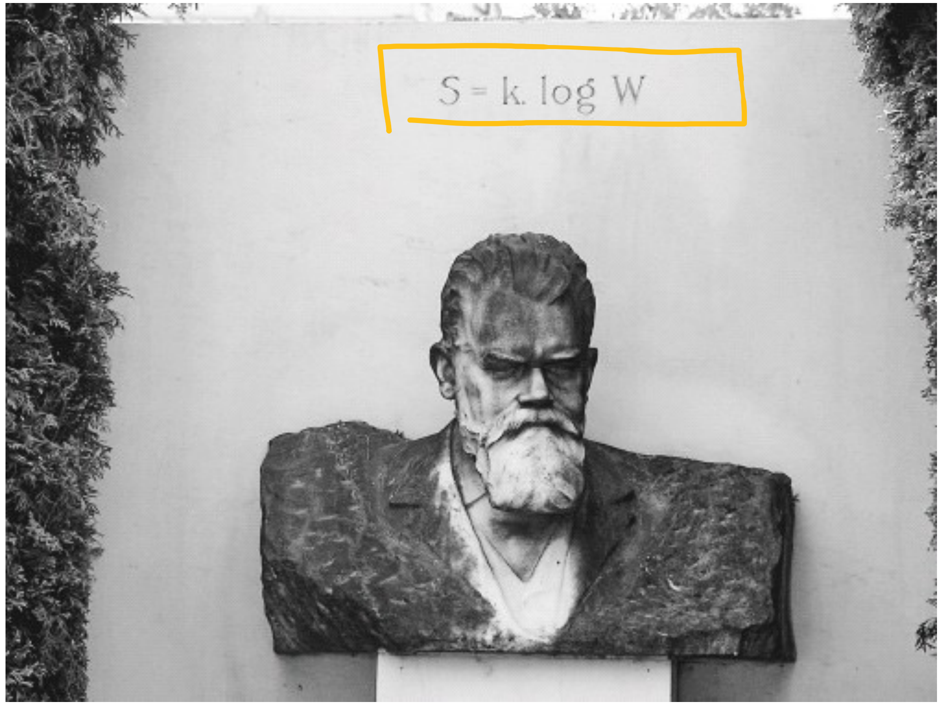


Macrostate: Fast and slow particles are completely mixed.

Second Law of Thermodynamics: In an isolated system, entropy will always increase until it reaches a maximum value.

Second Law of Thermodynamics (Statistical Mechanics Version): In an isolated system, the system will always progress to a macrostate that corresponds to the maximum number of microstates.

Boltzmann Entropy



Boltzmann's tomb, Vienna, Austria

The entropy S of a macrostate is k times the natural logarithm of the number W of microstates corresponding to that macrostate.

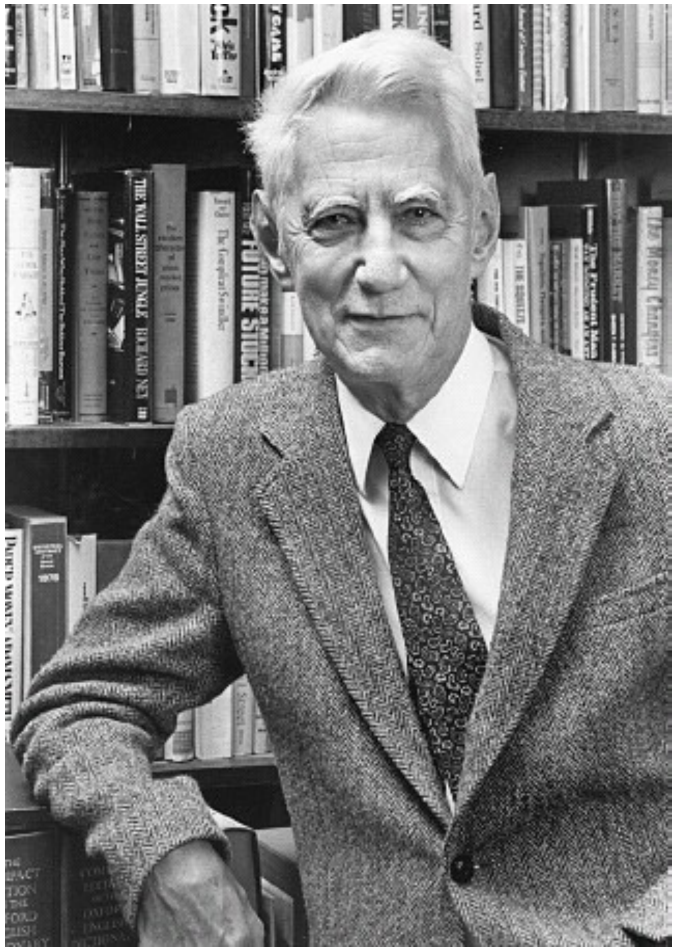
k is called "Boltzmann's constant". This constant and the logarithm are just for putting entropy into a particular units.

General idea: The more microstates that give rise to a macrostate, the more probable that macrostate is. Thus *high entropy = more probable macrostate.*

Second Law of Thermodynamics (Statistical Mechanics Version):

In an isolated system, the system will tend to progress to the **most probable macrostate**.

Shannon Information



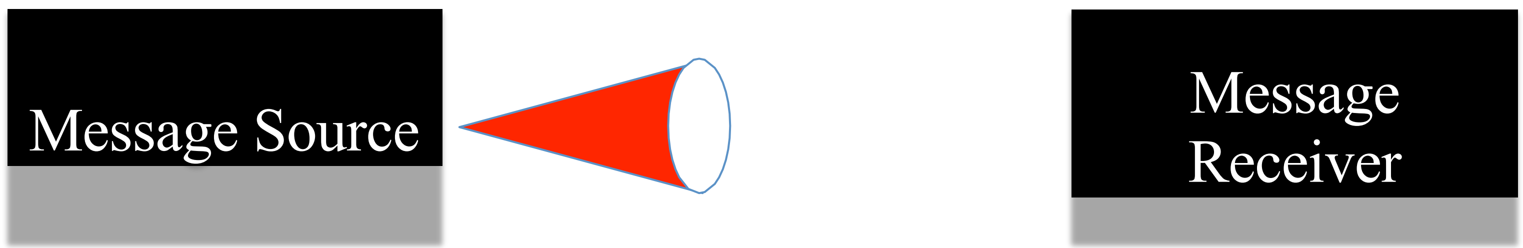
Claude Shannon, 1916-2001

Shannon worked at Bell Labs (part of AT&T)

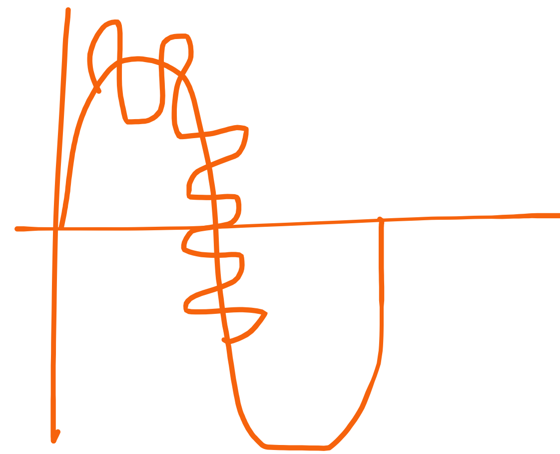
Major question for telephone communication: How to transmit signals most efficiently and effectively across telephone wires?

Shannon adapted Boltzmann’s statistical mechanics ideas to the field of communication.

Shannon's Formulation of Communication



Message (e.g., a word)

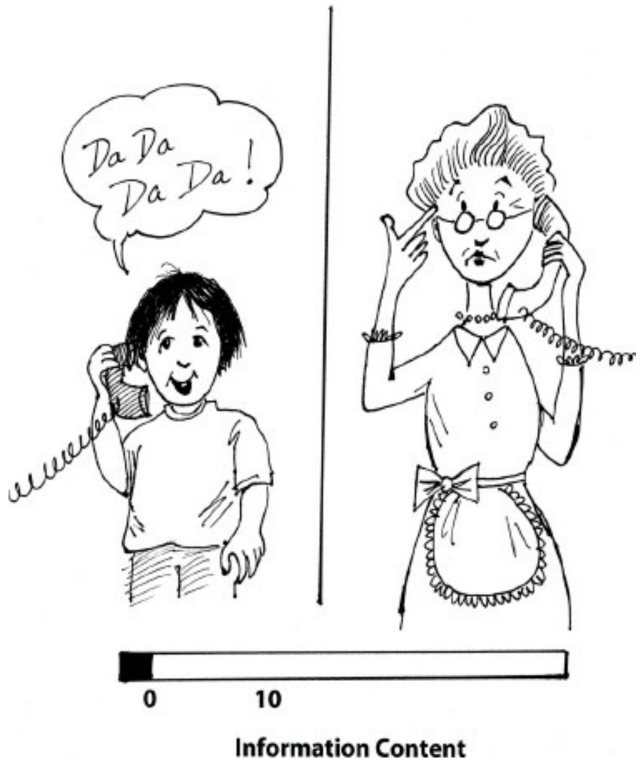


Message source : Set of all possible messages this source can send, each with its own probability of being sent next.

Message: E.g., symbol, number, or word

Information content *H* of the message source: A function of the number of possible messages, and their probabilities

Informally: The amount of "surprise" the receiver has upon receipt of each message

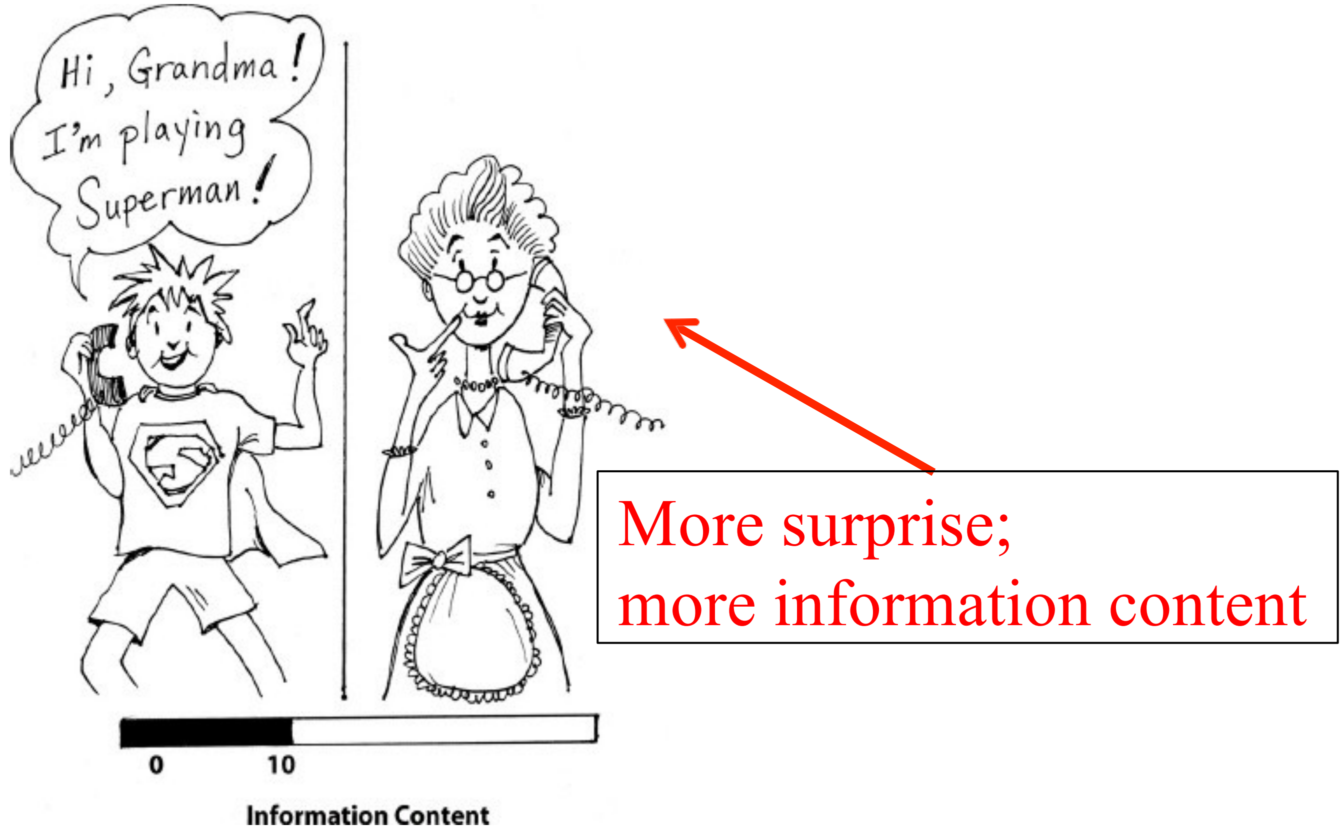


No surprise;
no information content

Message source: One-year-old

Messages:
"Da" Probability 1

InformationContent (one-year-old) = 0 bits



Message source: Three-year-old

Messages: 500 words $(w_1, w_2, \ldots, w_{500})$
Probabilities: $p_1, p_2, \ldots, p_{500}$

InformationContent (three-year-old) > 0 bits

How to compute Shannon information content

- Assume $k = 1$. Slot machine example. Calculate $S = \log W$ for two macrostates.

Boltzmann Entropy

Microstate: Detailed configuration of system components (e.g., “apple pear cherry”)

Macrostate: Collection of microstates (e.g., “all three the same” or “exactly one apple”)

Entropy S : Assumes all microstates are equally probable

ln

$$S(macrostate) = k \log W$$

where W is the number of microstates corresponding to the macrostate.
S is measured in units defined by k (often “Joules per Kelvin”)

Shannon Information

Message: **E.g.,** a symbol, number, or word.

Message source: A set of possible messages, with probabilities for sending each possible message

Information content H :
Let M be the number of possible messages. Assume all messages are equally probable.

$$H(message\,source) = \log_2 M$$

H is measured in “bits per message”

General formula for Shannon Information Content

Let M be the number of possible messages, and p_i be the probability of message i. Then

$$\boxed{H(message\,source) = -\sum_{i=1}^{M} p_i \log_2 p_i}$$

$$H(messagesource) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

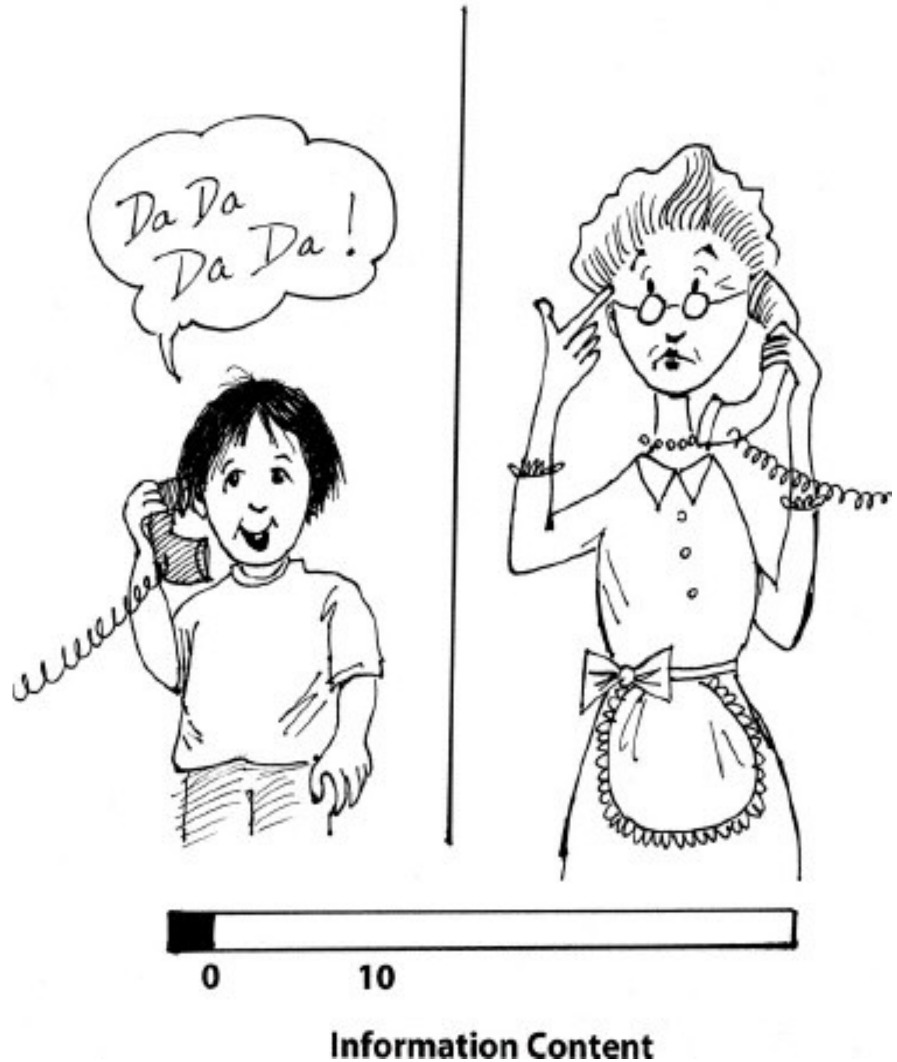
$$= -\sum_{i=1}^{N} \frac{1}{M} \log_2 \frac{1}{M}$$

$$= -\log_2 \frac{1}{M}$$

$$= -\log_2 M^{-1}$$

$$= \log_2 M$$

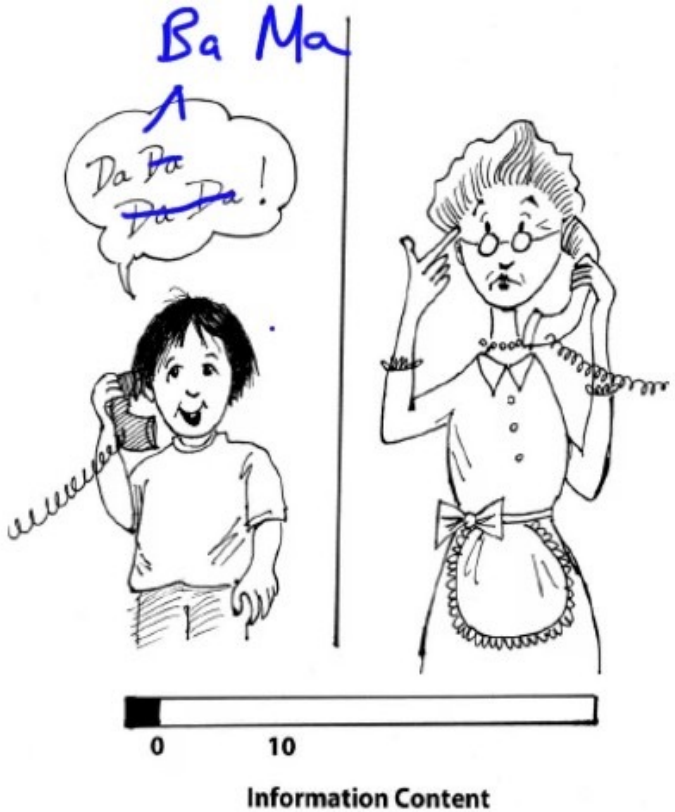
Message source: One-year-old: {“Da”, probability 1}



$$H(\text{one-year-old}) = -\sum_{i=1}^{N} p_i \log_2 p_i = 1 * \log_2 1 = 0 \text{ bits}$$

$M = 1$

$H(\text{one-year-old}) = \log_2 1 = 0$



$M = 3$

$H = \log_2 3 \simeq 1.58$

Message source: Fair coin:
("Heads", probability .5)
("Tails", probability .5)

$$H\,(\text{fair coin}) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

$$= -\left[\left(.5\log_2 .5\right) + \left(.5\log_2 .5\right)\right]$$

$$= -\left[.5(-1) + .5(-1)\right]$$

$$= 1\,\text{bit} \quad \text{(on average, per message)}$$

Message source: Biased coin:
("Heads", probability .6)
("Tails", probability .4)

$$H(\text{biased coin}) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

$$= -\left[(.6\log_2 .6) + (.4\log_2 .4)\right]$$

$$= .971 \text{ bits}$$ (on average, per message)

Message source: Fair die:
("1", probability 1/6)
("2", probability 1/6)
("3", probability 1/6)
("4", probability 1/6)
("5", probability 1/6)
("6", probability 1/6)

$$H(\text{fair die}) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

$$= -6\left(\frac{1}{6}\log_2 \frac{1}{6}\right)$$

$$\approx 2.58 \text{ bits (per message, on average)}$$

Text Analysis

- Text info content: One way to do text analysis

Info content of text [one way to measure it]: based on relative frequency of word in the text. Roughly measures *compressibility*

E.g., "to be or not to be"

to: 2 $\frac{1}{L}$	relative frequency: 2/6
be: 2	2/6
or: 1	1/6
not: 1	1/6
Total words: 6	

$$H(\text{"to be or not to be"})$$

$$= -\left[\left(\frac{2}{6}\log_2\frac{2}{6}\right)+\left(\frac{2}{6}\log_2\frac{2}{6}\right)+\left(\frac{1}{6}\log_2\frac{1}{6}\right)+\left(\frac{1}{6}\log_2\frac{1}{6}\right)\right]$$

$$\approx 1.92$$

More generally: Information content = average number of bits it takes to encode a message from a given message source, given an "optimal coding".

H("to be or not to be") $\approx$ 1.92 bits per word on average

This gives the compressibility of a text.

See "Huffman Coding".

Shannon Information Content versus Meaning

Shannon information content does not capture the notion of the *function* or *meaning* of information.

The *meaning* of information comes from **information processing**.