Many faces of information: irreversibility, ignorance, inference

Rajaram Nityananda
Azim Premji University
Bengaluru 560100
http://www.ncra.tifr.res.in:8081/~rajaram
rajaram.nityananda@gmail.com

The best evidence that the subject has reached wide significance A best-selling paperback by James Gleick available in all airports.

Recommended!

Measuring information; bit by bit

- Information is the removal of uncertainty / ignorance. So we need to measure ignorance
- W, the number of equally probable options (before receiving the information) is a reasonable measure so is $p=1/W$
- W multiplies, log W adds (can also use -log p)
- 2 is a good base to use
- Generalises to $S=-\sum p_i \log_2 p_i$

W=2

W=64

Real message have correlations

- Symbols are not independent e.g q is almost always followed by u
- Shannon entropy can be generalised to take care of this, just use the joint probability distribution for long messages
- With proper coding, the message can be compressed. This happens all the time nowadays e.g gzip

Channel Capacity

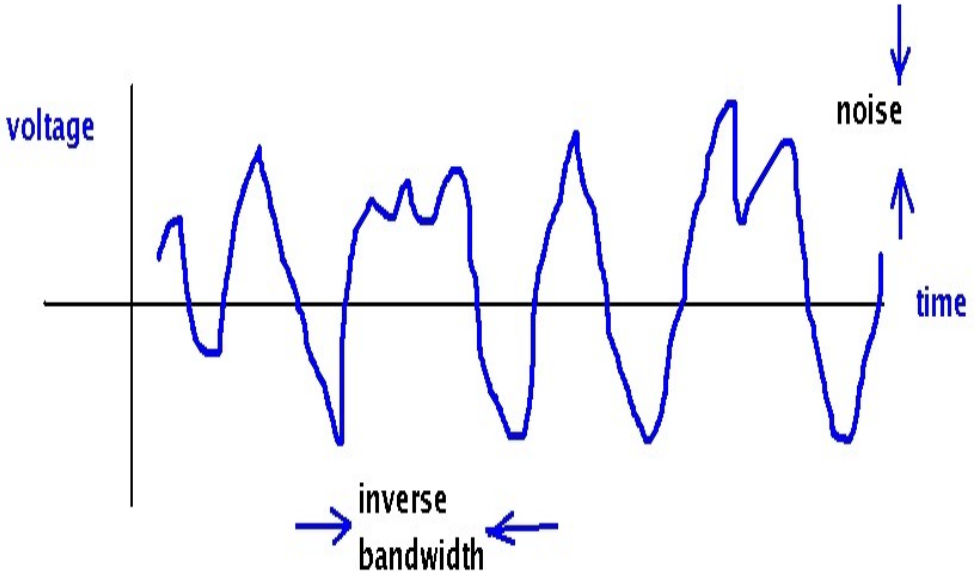
Log of no. of possible messages to the base 2

there are (26)^4 entries in a dictionary of four letter words, all the way from AAAA to ZZZZ

$$S = 4 \times \log_2 26 \approx 18.8 \ll 26 \times \log_2 4 = 52$$

Number of symbols in the alphabet is $\sim$ S/N

The number of letters in the message is 2Bt



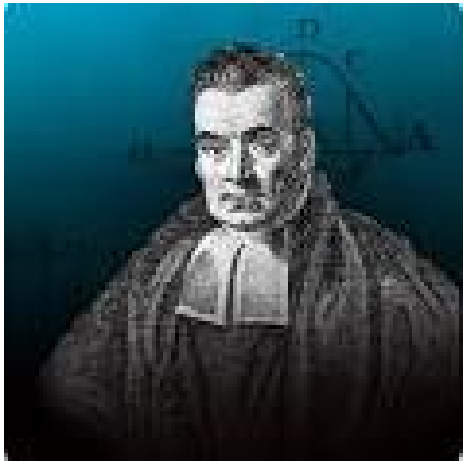
Statistics, The Big Fight

- Error on the mean

$$\bar{x} = \frac{(x_1 + x_2 + ... + x_N)}{N} \pm \frac{\sigma}{\sqrt{(N)}}$$

- The usual calculation assumes we take ***many*** samples from an ***known*** underlying distribution and works out how they are distributed – this is a deductive process (F for Fisher or frequentist)
- In real life, you have ***one*** sample, and you have to give error bars on the parameter (say mean) of the ***unknown*** underlying distribution!

Problems of inference / induction / inversion



Bayes

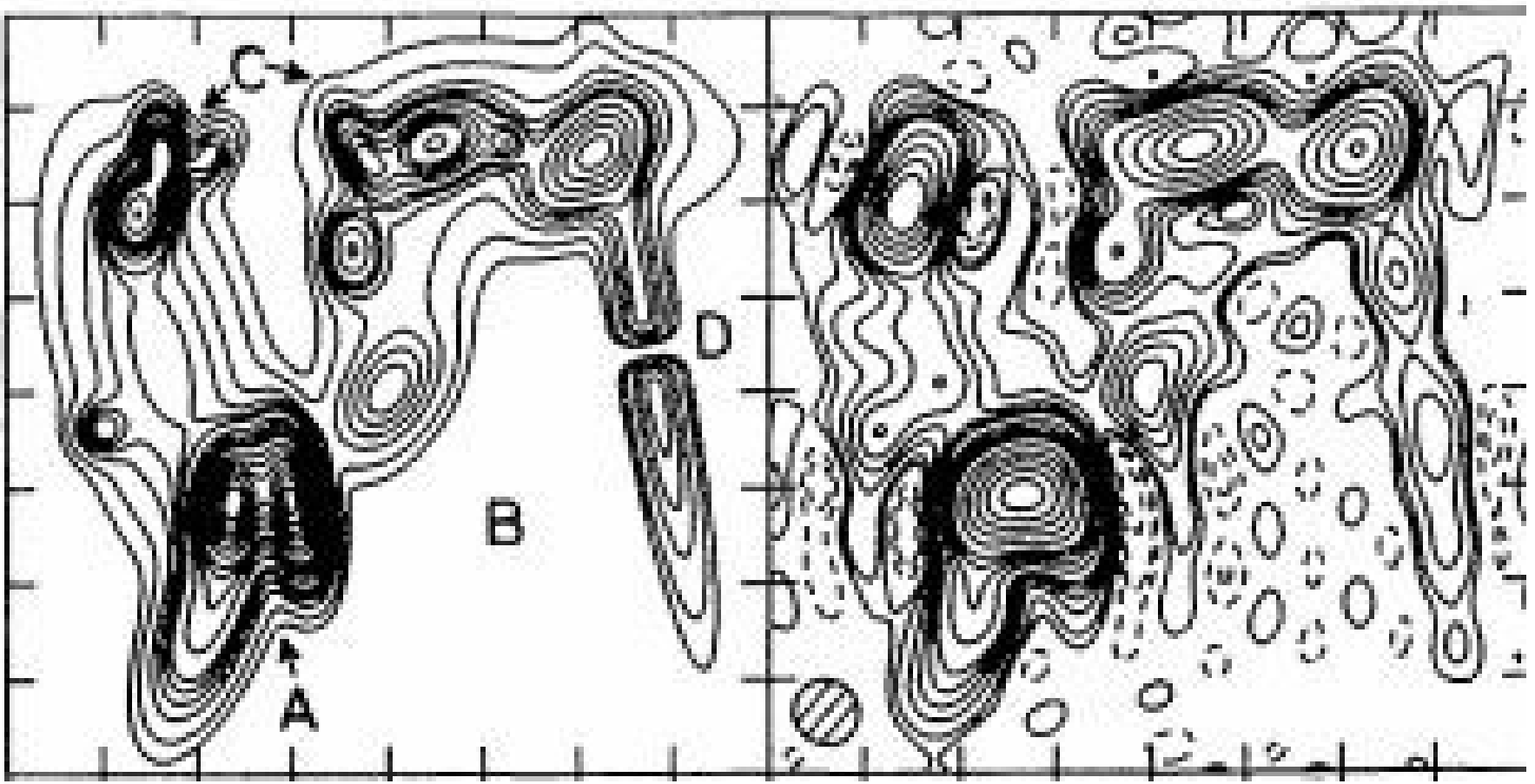
Laplace

$$P(H|D)=\frac{P(D|H)P(H)}{P(D)}$$

Jeffreys

Jaynes

Maximum entropy image restoration
A Bayesian strategy?



Ramesh Narayan and RN, ARAA 1986

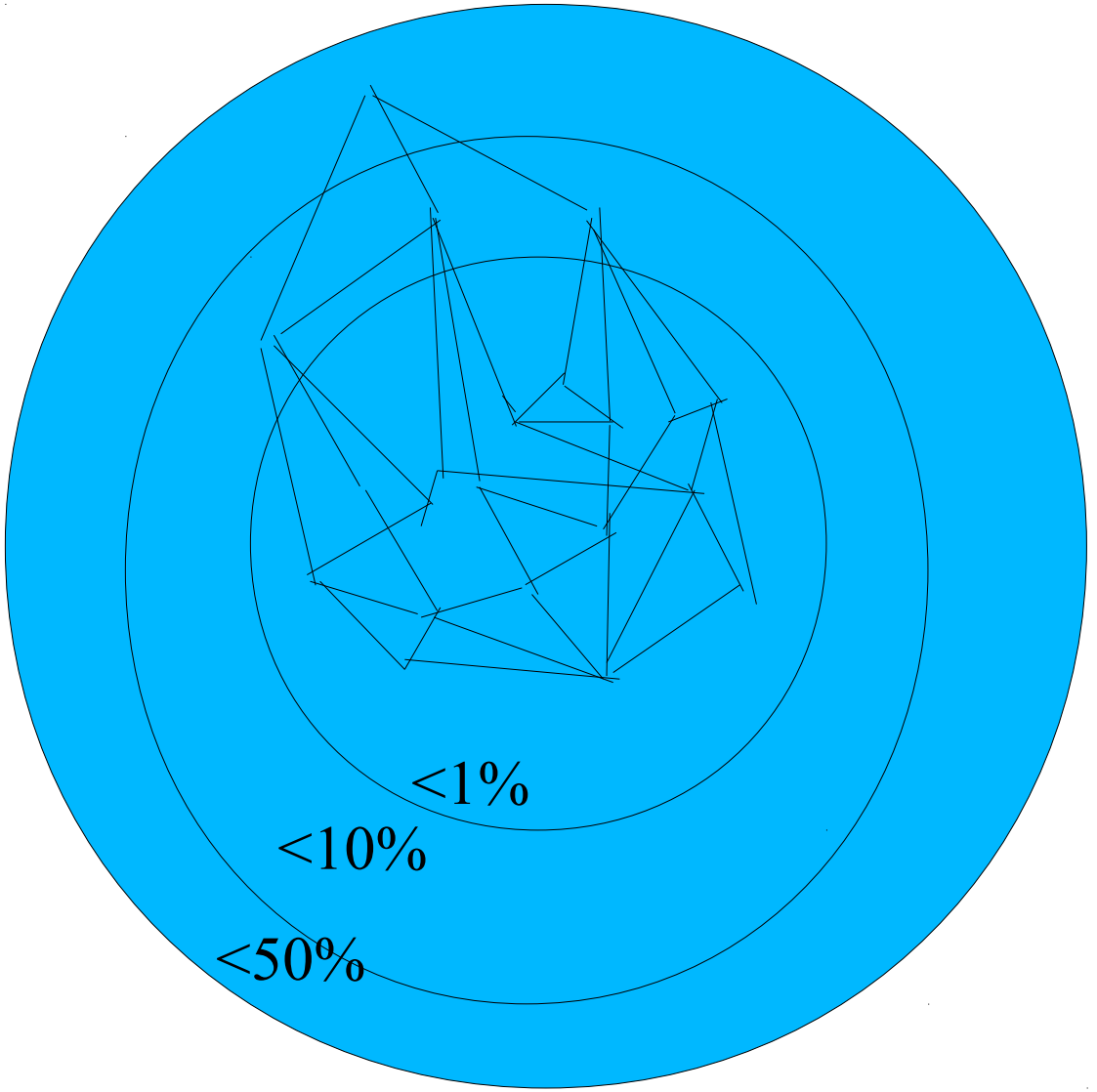
The Prior

- *P(H)* is required by Bayes theorem
- How can we say anything about *H* without data? BUT
- Maybe those who do without a prior are just hiding it somewhere
- "Entropy" is now the logarithm of the prior
- Stating the prior is an honest exhibition of our assumptions and sensitivity to the prior is an honest appraisal of the uncertainty of our methods

Irreversibility

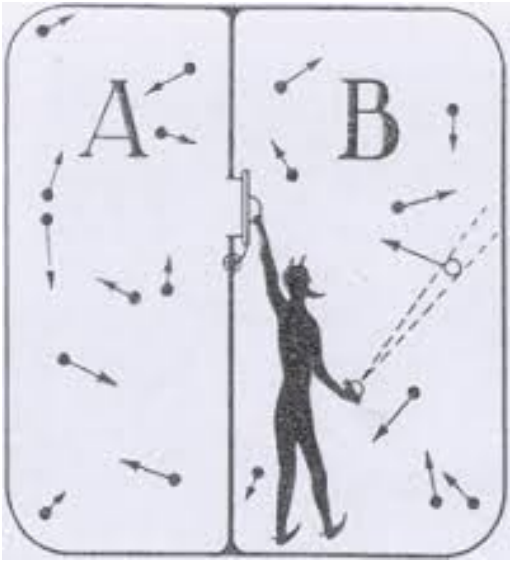
- Boltzmann gave the same *-p ln p* type of formula for entropy, starting with *ln W*
- 'Default' is for the system to be in a "state" with a large value of *W*. (Equlilibrium)
- There is no mysterious force driving a system irreversibly towards equilibrium! Microscopic reversibility prevails. Equilibrium is just tyranny of the majority
- Need 'molecular chaos', i.e there should be no constraint preventing the system from exploring all the states

Schematic view of equilibrium and fluctuations



Information can be traded for entropy

- Maxwell demon -separates hot and cold



- Szilard exorcism: you need information on molecular speeds
- Destroying information increases entropy (e.g erasing the results of computation). reversible and irreversible computation (Bennet, Landauer of IBM)

Information at the foundations of physics

- Information is physical
- Quantum information, communication, computation
- Black hole entropy and “information loss”
- Is all physics about information?
- Qbism – information is always relative to someone (one?….who ? What ? ….)
- Nagarjuna? “nothing exists by itself, only in relation to other things”

jambhulkars96@yahoo.in
+91-9967320933

Claude Elwood Shannon: father of information theory

MS thesis

PhD

War work

Bell Labs

MTC

aftermath