The Edge of Knowledge

H.J. Pirner

Institut für Theoretische Physik/Marsilius Kolleg Heidelberg

Our educational background:

- We learn:

- to be precise,
- define exactly,
- come to certain answers,
- decide rationally,
- determine well

- We push away:

- Vague concepts
- Indefinite ideas
- Uncertainties

- Undecidedness
- Indetermination

But this does not work:

- Uncertainties and indeterminacies creep upon us, even if we do not like them
- Especially at the frontier of science we have to deal with them
- With increasing knowlege (doubling every 20 years) also the edges of knowlege will grow accordingly

Outline:

- I. Five characteristic uncertainties

  Unreliable memory, indeterminate experiment, uncertain prognostics, vague statements, Plato‘s hint

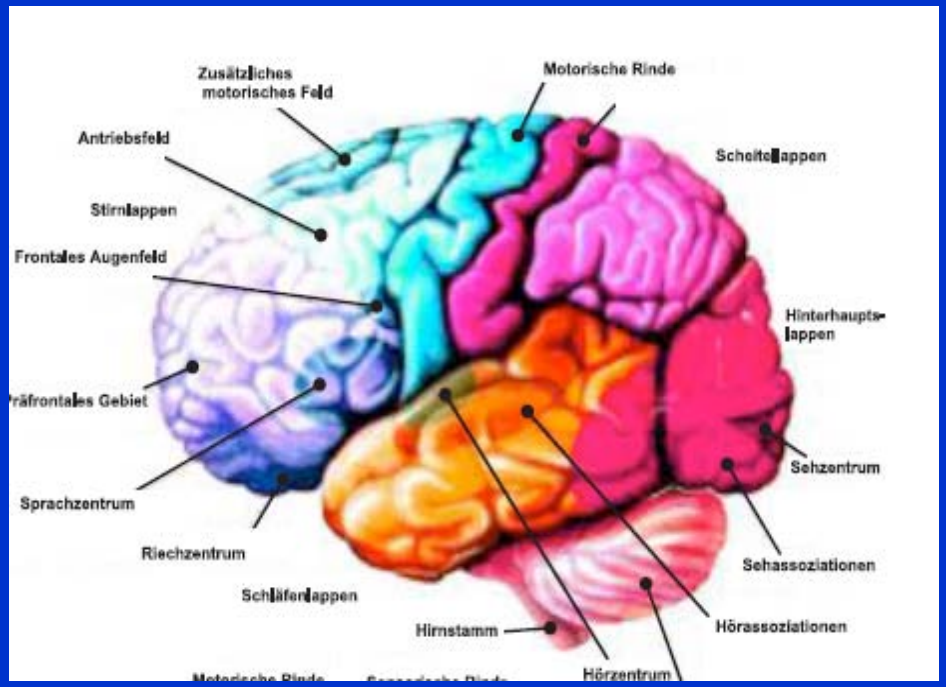
- II. Get more information !

  Information Theory, Inference and the Value of Information

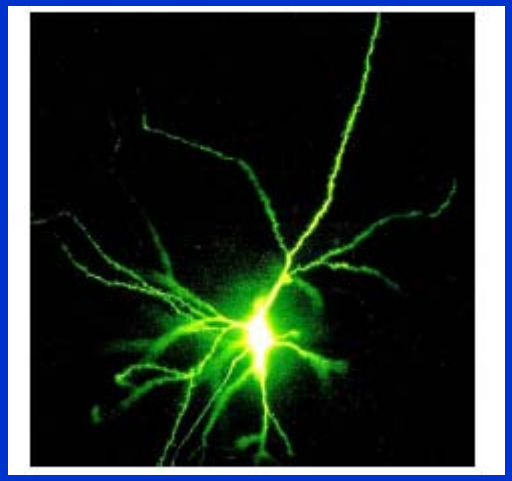
- III. Metaphors, signs or games help?

  Language, Philosophy and Economics

1. Unreliable Memory :



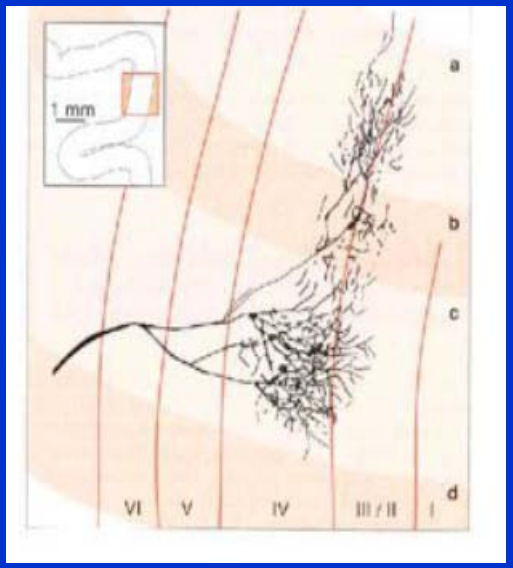
Cortex contains 10 000 million neurons
Each neuron contacts 10 000 others
Dendrites give the input
long axon transports the signal (mV)



Enhancement of Synaptic Pattern

17 day old cat

Adult cat



Single Neuron (Perceptron):

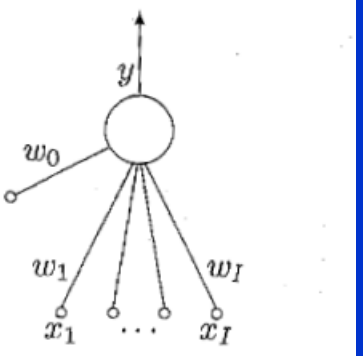


Figure 39.1. A single neuron

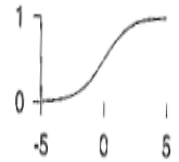
Aktivation of the neuron:

$$a = \sum_i w_i x_i,$$

Output function y(a):

(logistic function).

$$y(a) = \frac{1}{1+e^{-a}} \qquad (y \in (0,1)). \qquad (39.3)$$

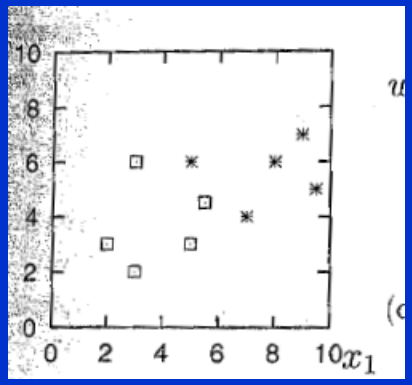


Input potentials x(i) und Synaptic weights w(i) produce Activation a. Output potential y(a) defines whether neuron fires or does not fire

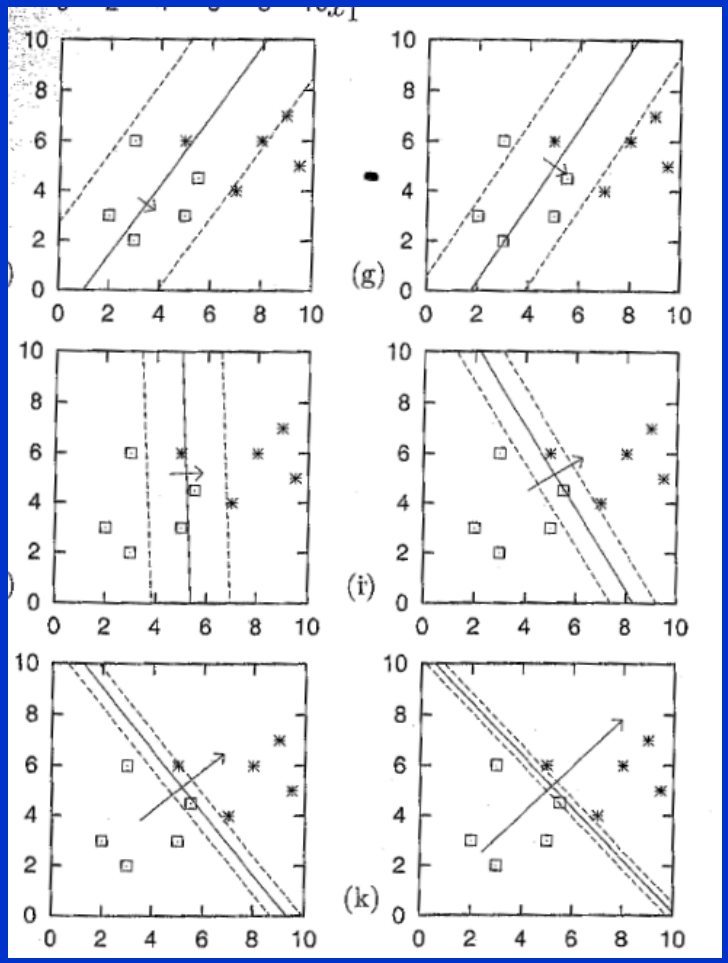
.

Learning:

Stars (t=0)and squares (t=1):
Separate them!



-→Evolution of Learning =Adjustment of synaptic weights w1 and w2, to minimise error= t-y(x1,x2,w1,w2), δw(i)= η error x(i) manipulate line of separation between stars and squares
30,80,500,3000,10 000, 40 000 Iterations

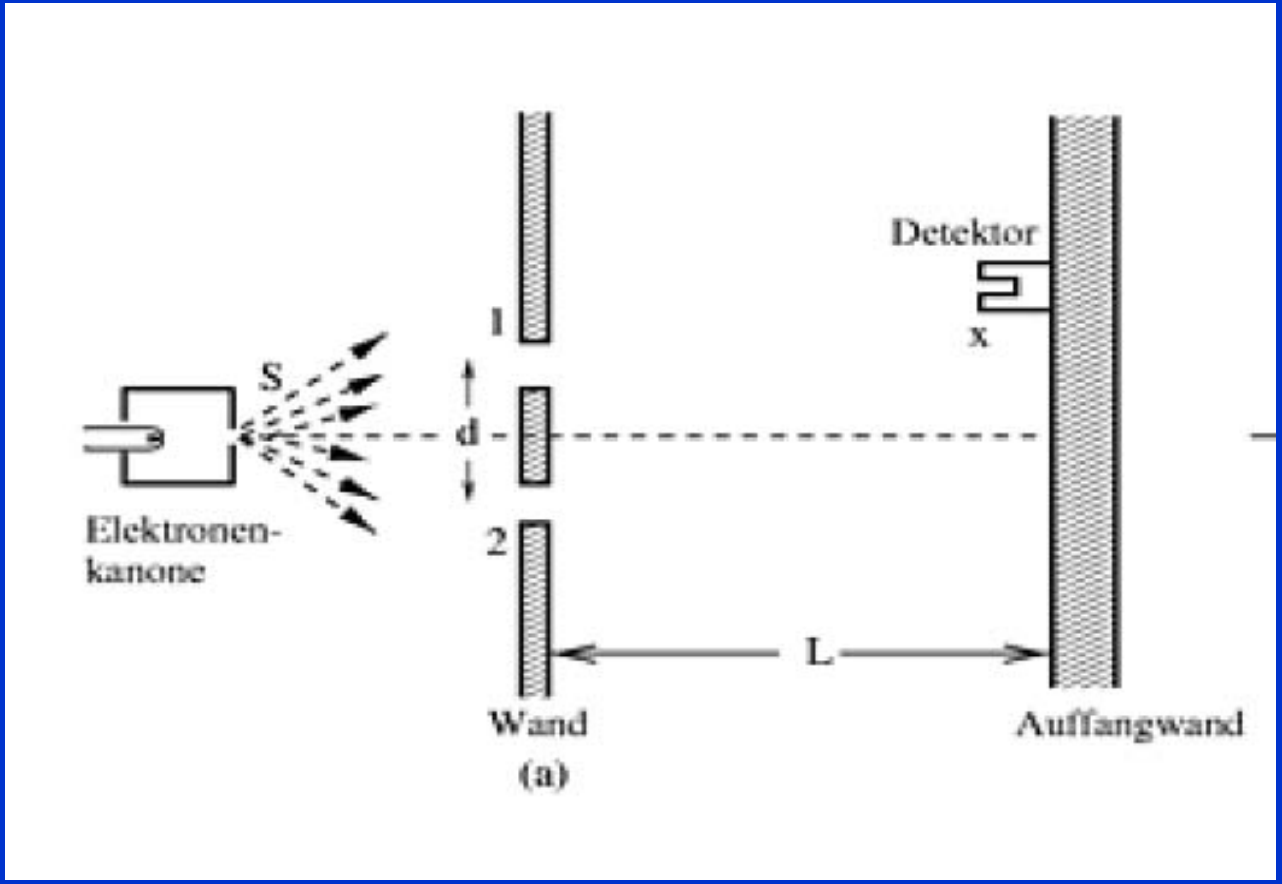


„Bad“ memory?

- Problem may not be separable.
- Need change of perspective.
- More than one minimum exists.
- Different answers alternate
- Synaptic weights age?
- Memorizing may change the neural net

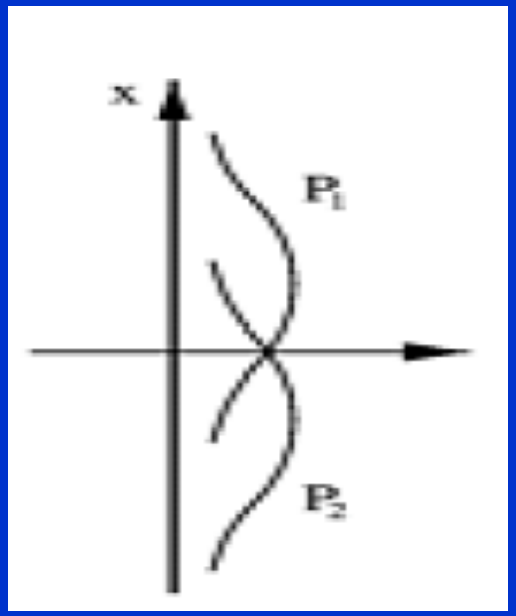
2. Indeterminacy in Quantum Physics

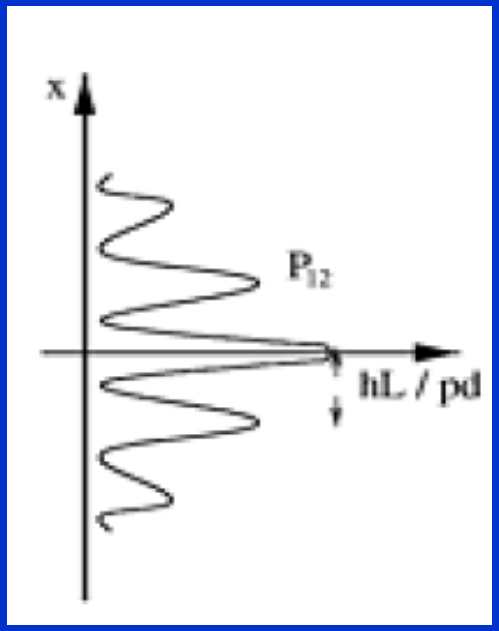
The Double Slit Experiment



Classical versus Quantum:

Violation of the Aristotelian logic „tertium non datur“ for probabilities(!)





Upper Slit open: p1
Lower Slit open:p2
Both slits open: p1+p2

Both slits open: Interference Pattern
We do not know through which slit the electron went

3. Uncertain Prognostics:

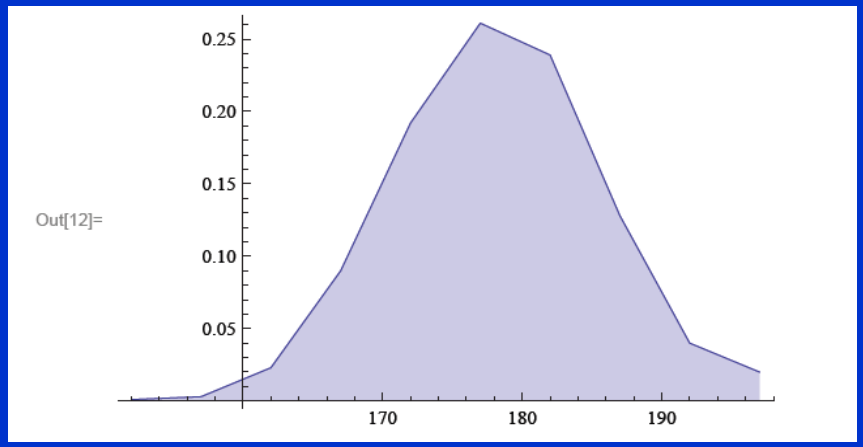
- Meet a guy. Can you estimate the height of the guy? How accurately?
- Meet a banker. Can you estimate the income of the guy? How accurately.

When our information about a system are limited, then we can make predictions, based on probabilities. Differentiate between mediocristan and extremistan.

Predictions in Extremistan cannot give an error:

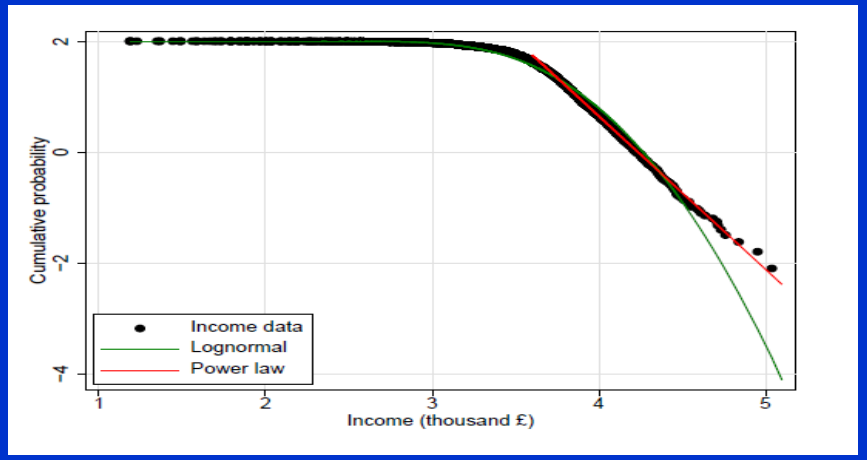
Mediocristan

Height of German males
relative
number/height(cm)
Bell shaped curve (normal distribution)



Extremistan

Cumulative income distribution.
Log -normal distribution, but the highest incomes obey a power law with index α=1.4,



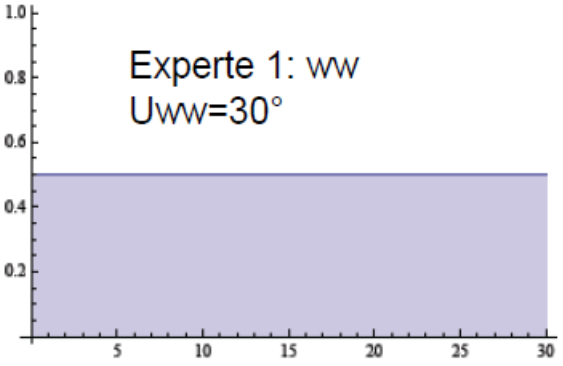
Marsilius 2011

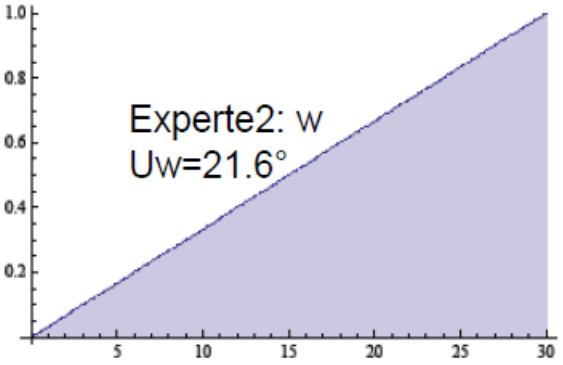
Power law allows to estimate the mean, but not the variance

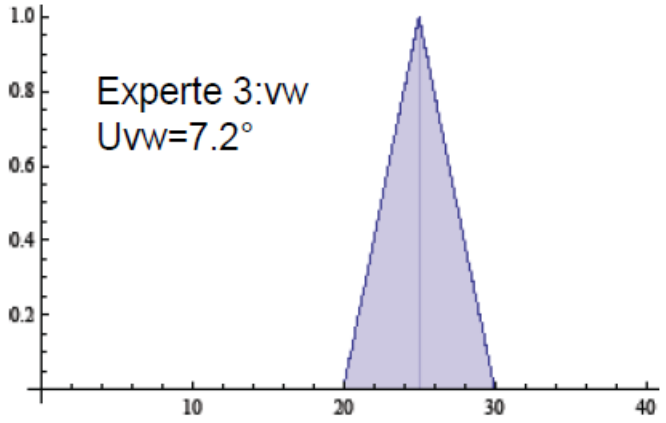
4. Vagueness:

- „This room is warm“, „the color is red“,etc
- Vague statements like these are abundant in our language
- Philosophers do not like them, vague sentences are neither wrong nor true
- Conclusions based on them are unclear
- Engineers propose a „Fuzzy Logic“ with gradual „membership functions“ m:

Vague Statement:
„ This room is warm“







Mathematische Definition der Unbestimmtheit

$$U(A) = -\int dt\,(m(t)\,\mathrm{Log}[2,m(t)] + (1-m(t)\,\mathrm{Log}[2,1-m(t)])$$

Membership function m=1 (m=0) means the statements is correct (incorrect)

Realm of Fuzzy Logic

- One can define a
- Fuzzyness:

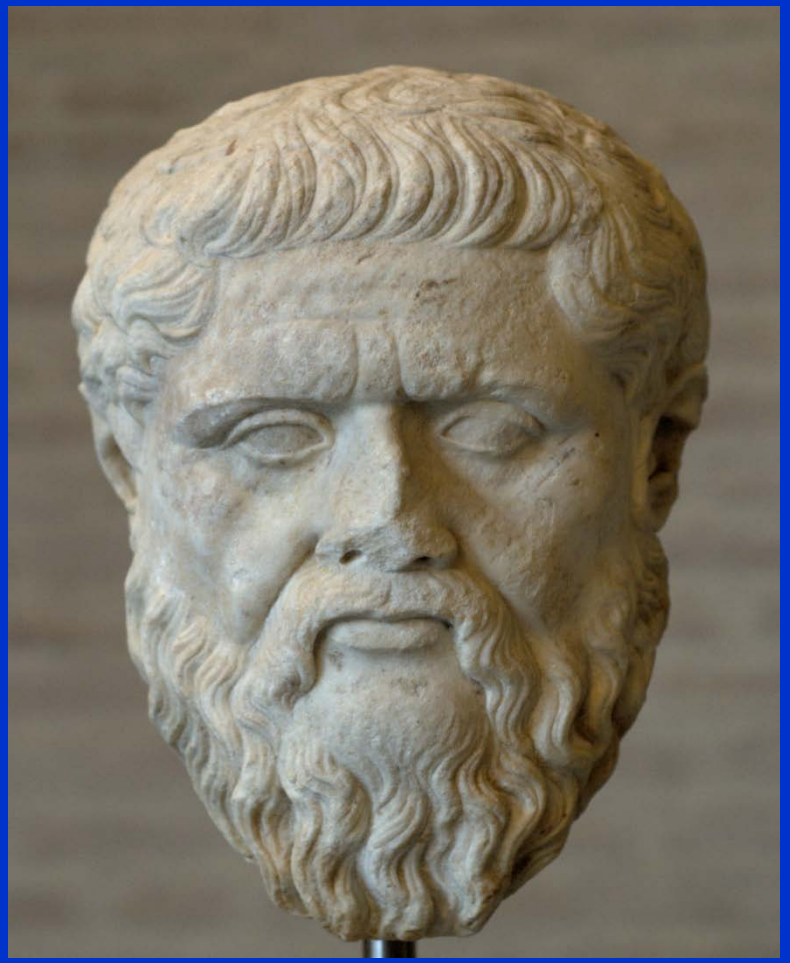
$$U(A) = -\int dt\,(m(t)\,\mathrm{Log}[2,m(t)] + (1-m(t)\,\mathrm{Log}[2,1-m(t)])$$

- One can handle „logical“ operations e.g.
- $m(A \cap B) = \min[m(A), m(B)]$

But there also some paradoxa:
$m(A \cap \neg A) = 0.5$ when $m(A) = 0.5$,
But we would expect that the above statement has m=0.
Therefore we cannot give truth functionality to the membership functions, they are more like opinions
This method has been extensively used in control engineering:
Smooth running of subways,
Control of cameras by symbols etc.

5. Plato

Plato (428/427 BC[a] – 348/347 BC),
was a Classical Greek philosopher,
mathematician, student of Socrates,
writer of philosophical dialogues,
e.g. Philebos,
founder of the Academy in Athens,
the first institution of higher learning
in the Western world. Along with his mentor,
Socrates, and his student, Aristotle,
Plato helped to lay the foundations of
Western philosophy and science



Plato‘s Hint:

Plato‘s Hint:

Soc. Let us be very careful in laying the foundation. Let us divide all existing things into two, or rather, if you do not object, into three classes. Were we not saying that God revealed a finite element of existence, and also an infinite? Let us assume these two principles, and also a third, which is compounded out of them; but I fear that am ridiculously clumsy at these processes of division and enumeration. I say that a fourth class is still wanted. Find the cause of the third or compound, and add this as a fourth class to the three others. Let us begin

The Indefinite (infinite), the definite (finite), the mixture of the two, and the cause for their togetherness are the basic elements of the world.

II. Information as Antagonist to Uncertainty:

If you are uncertain, more information can help you.

- (Potential) Information has been quantified by Shannon:
- Take a chain of letters: ABBBBA….
- The probability for p(A) =2/6 for p(B) = 4/6
- The surprise effect to get a letter A transmitted is higher than with the letter B
- For two independent events p=p1*p2
- I(p1*p2)= I(p1)+I(p2)

Shannon‘s Information

$$I(p) = Log_2\left(\frac{1}{p}\right) = -Log_2\,(p)$$

$$I\left(\frac{1}{2}\right) = 1\ (\text{Bit}).$$

- This definition fulfills the expectations.
- I bigger, if p smaller
- I for independent events
- Unit 1 Bit for yes/no

$$H = < I(p) > = \sum p_i\ I(p_i) = -\sum p_i\, Log_2\,(p_i)$$

Mean information

Inference with Bayes Statistics

$$P(y|x) = \frac{P(x|y)P(y)}{\sum_{y'} P(x|y')P(y')}$$

Relation between conditional probabilities

Example: *Charles is tested for swine-flu. He belongs to a group in which the risk to get the sickness is 1%. The test has a certainty of 95%. The test indicates that Charles has the swine flu*. What is the probability that he has the swine flue?

10 000 = number in his risk group

100 (1%) have s-f 9900 do not have sf

Test: 95 test infected , 5 not infected ---------- 495 test infected, 9405 not inf.

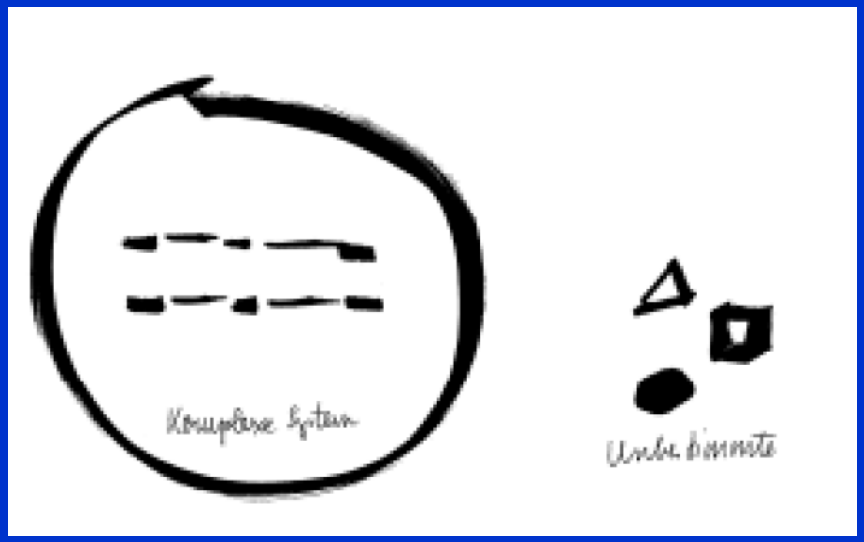
P (Charles to have s.f.)= 95/ (95 +495)= 16%

Quality of information

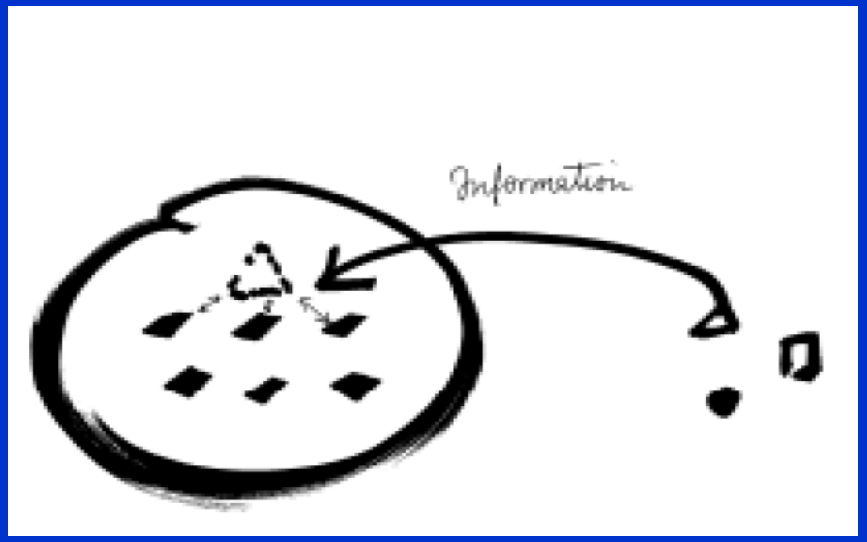
- In Infodynamics one handles information, like energy in thermodynamics
- Energy is conserved, but one can qualify energy by the amount of entropy contained in it. F= E – T S. A lot of entropy means not much free energy to do work.
- My idea is to use „fuzzyness“ to qualify information

System Theory:

Separate the definite system and the indefinite environment:

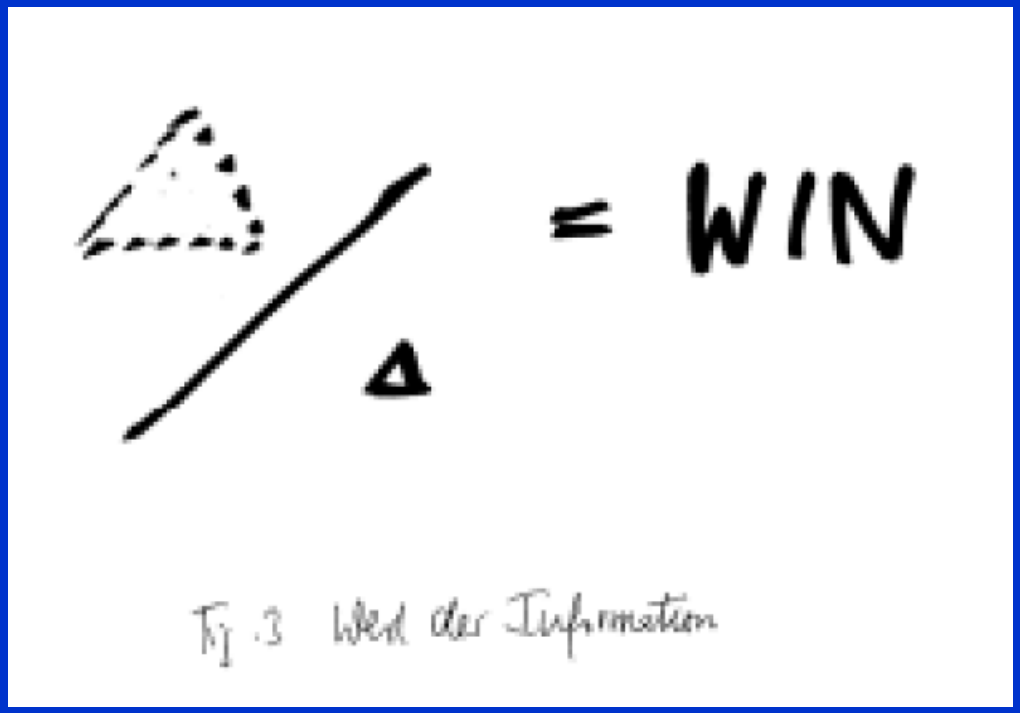


Process of Information:



Change in complexity in the system and
a change in the indefiniteness of the environement

Value of Information



Value of Information = Δ Complexity / Δ Log [Fuzzyness U]

$$U(A) = -\int dt\,(m(t)\,\mathrm{Log}[2,m(t)] + (1-m(t)\,\mathrm{Log}[2,1-m(t)])$$

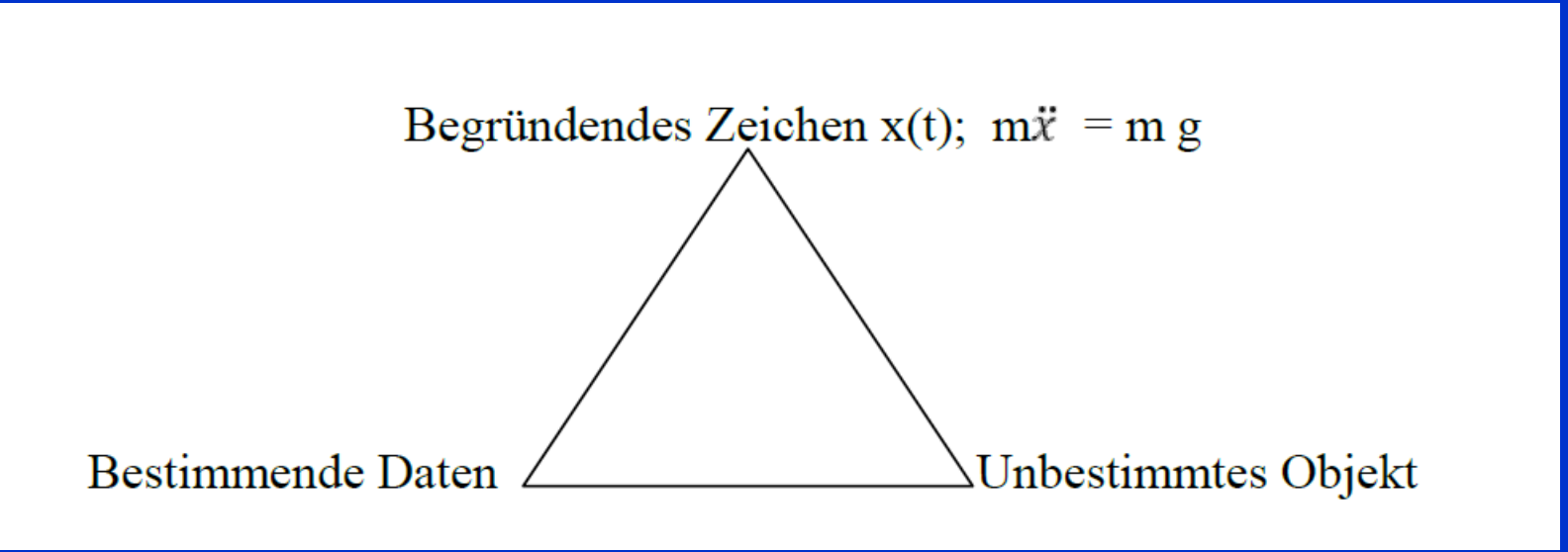
III. How to determine the Undetermined

- With metaphores, also in science
  (Dark matter, dark energy, inflaton…)
- The metaphore carries the undetermined into another field, where some knowledge already exists. (This transfer is also used in psychoanalysis)
- Physicists like to make their theory „beautiful“, i.e. transport it into the realm of „art“, but this does not always make the theory correct.

Signs:

Sign contains three parts in semiotics: recall Plato.I use the example of the falling apple in mechanics

The reason for their togetherness(Idealism)



Determining data
„Phenomenology“

Undetermined Object
„Realism“

Practical Sciences:

- Human decision has to define first an aim, to be reached. (Differentiate descriptive and normative decision theory)
- Decisions for alternatives a(i) under uncertainty use probabilities p(j)
- The gain for alternative a(i) and p(j) is put into the gain matrix money (i,j)

Probability for rain p1=0.8/ no rain p2=0.2

Gain Matrix :

$$Money_{i\,j} = \begin{pmatrix} 10 & 10 \\ 15 & 8 \end{pmatrix}$$

For alternative 1 (plant rye) and rain 10 $ and also with no rain 10 $.

- With alternative 2 (plant wheat) one gains for rain 15 $ and with no rain 8 $

Expectation for the gain:

$$\text{Alternative } a_1 >= \; Money_{11} * p_1 + Money_{12} * p_2 = 10\,€$$

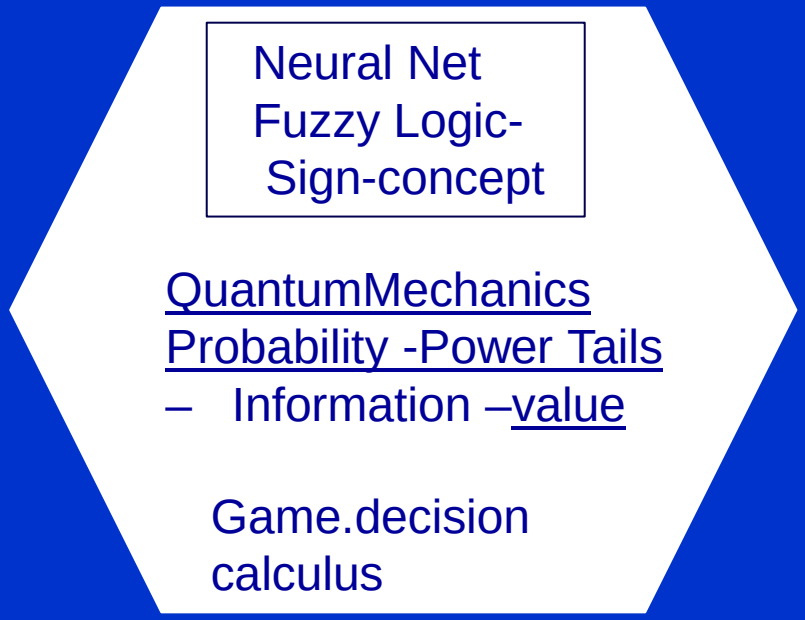
$$\text{Alternative } a_2 >= \; Money_{21} * p_1 + Money_{22} * p_2 = 13.6\,€$$

Decision Theory

- If the situation is happening repeatedly, then the expectation value of the gain is a good measure for decision.
- But it may well be, that you prefer alternative 1, if you do not like to risk anything. (One has to define a personal risk/ profit curve, which parametrizes the profit as a function of the risk)
- More complicated is the situation when you have other persons‘ actions to take into account→ Game theory

Summary: Which Uncertainties? How to handle them?

Undefined, Vague



Uncertain

Undetermined

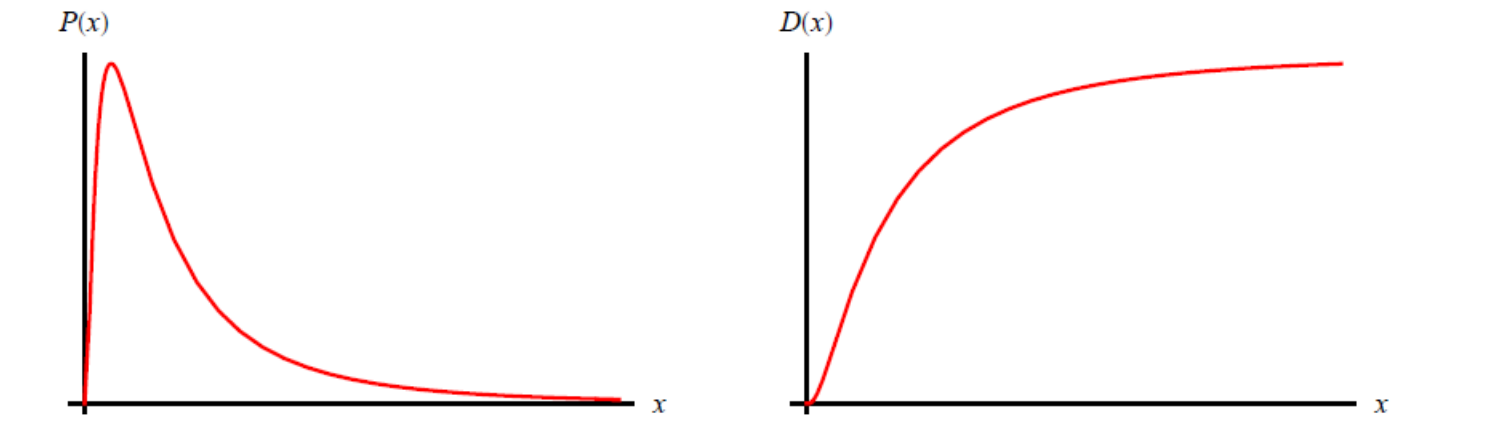
Undecided

Tao in Chinese philosophy

- „The Tao is something blurred and indistinct,
  how indistinct , how blurred
  Yet within it are images
  Yet within it are things
  How dim and how confused
  Yet within it is mental power
  because this power is most true
  Within it there is confidence“

Laotse , (600 B. C) Tao te ching, 道德經

Log Normal Distribution and Cumulative Distribution



Domain	Interval[{0, ∞}]
$P(x)$	$\frac{e^{-\frac{(M-\ln x)^2}{2S^2}}}{\sqrt{2\pi}\,Sx}$
$D(x)$	$\frac{1}{2}\left(\text{erf}\left(\frac{\ln x - M}{\sqrt{2}\,S}\right)+1\right)$

Methoden um Gehirnaktivität zu messen:

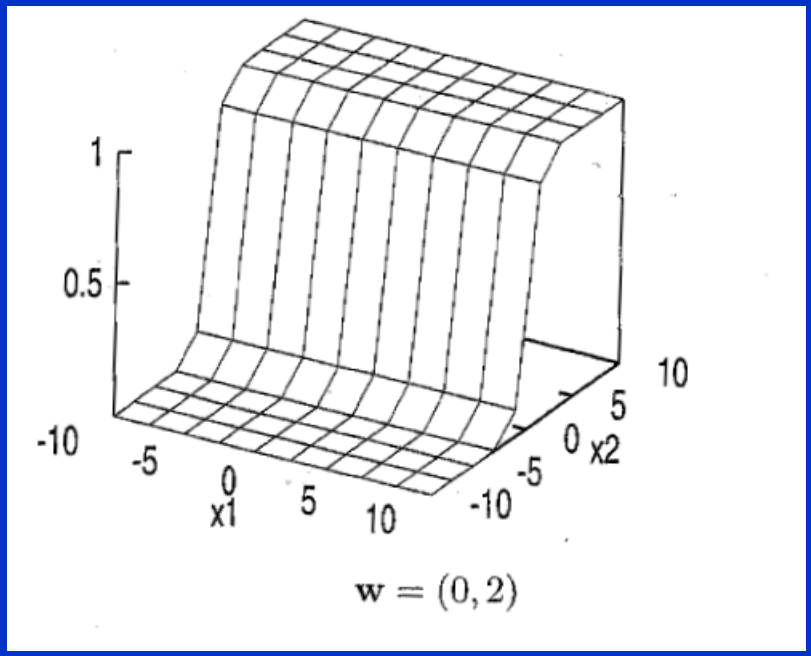
- Magnetoencephalographie misst extrem geringe(0.00000001*Erdfeld) Magnetfeldänderungen
- Positronen Emissions Tomographie ($e+e \rightarrow \gamma\gamma$) mit Positronen von O15 misst erhöhte Sauerstoff-konzentration
- Nukleare Magnetische Resonanz misst die Präzession des Kernspins im Magnetfeld, welche von der Wechselwirkung mit der Umgebung abhängt (z.B. Oxydationsgrad von Hämoglobin)
- Zwei Photonen Fluoreszensspektroskopie misst Ca
- Elektroden messen direkt die elektrischen Potentiale

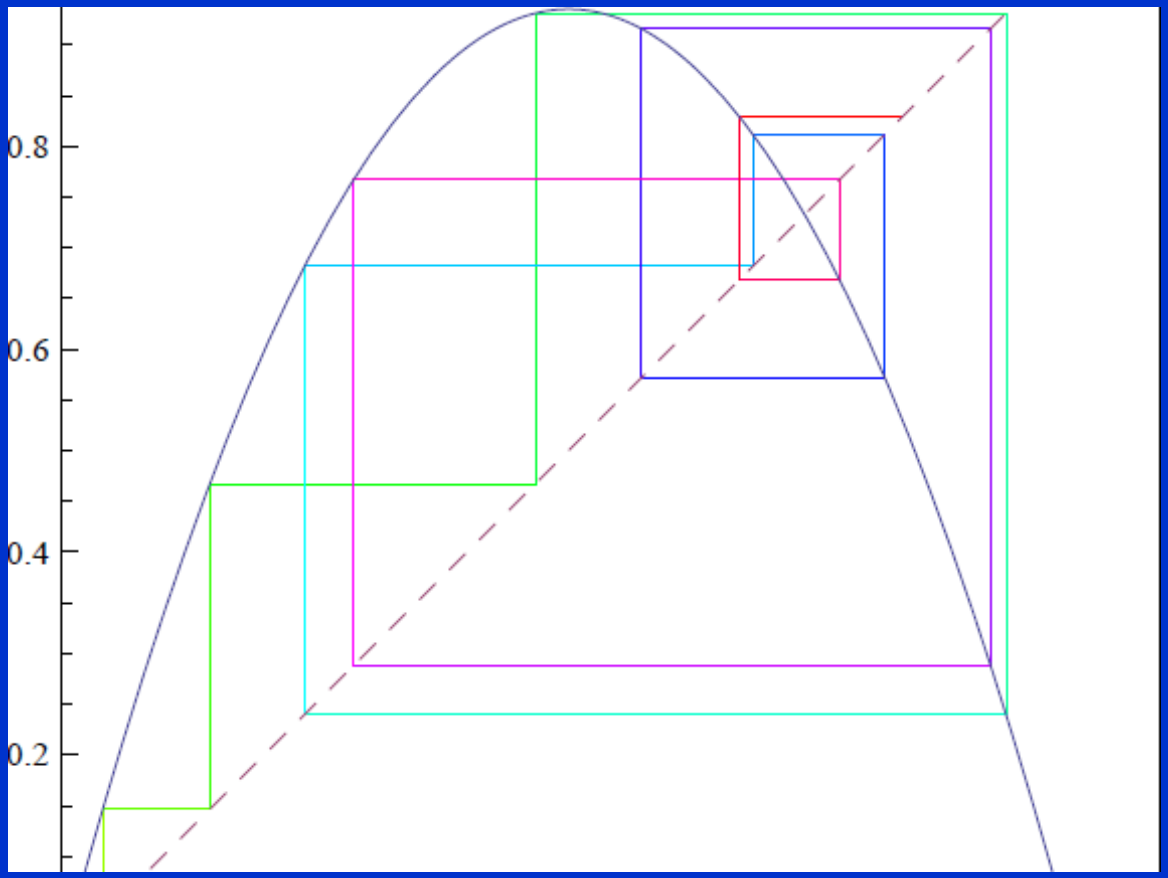
Aktivierungs-Funktion

y(a=x1 w1 +x2 w2) heisst A-Funktion. Zur Vereinfachung nehmen wir nur zwei Eingabedaten x1 und x2 und die zugehörigen Gewichter w1 und w2.

Die Aktivierungsfunktion ist links als Funktion von x1 und x2 dargestellt für w1=0 und w2=2

$$y(\mathbf{x};\mathbf{w}) = \frac{1}{1+e^{-(w_1x_1+w_2x_2)}}.$$




0.8
0.6
0.4
0.2