Reality as Simplicity

June 2009

Giulio Ruffini

Starlab®

Starlab®

Overview

Goal: To illustrate the importance of simplicity in various fields. To pose the question “why simplicity?”, and some hints.

1. Cognition from Information
2. A detour on computation: some questions
3. Simplicity
4. Why is simplicity important?
5. Fundamental neuroscience
6. The organism as a computer
7. Conclusions

Starlab®

Cognition, information, computation

Starlab®

Cognition from information

- Overview:
  - Brains (later organisms) receive information from the environment (via sensors), process it and transmit some information out to effectors (e.g., our hands) to control sensorial systems and to “act” on the environment thorough our bodies or BCIs.
  - Brains are computers. Information comes in, information goes out. Both are very important!
  - NB: the environment here includes the body, but some computation carried out by our bodies as well
  - What we call “reality” is a mental construct, a model. Simplicity is part of our modeling strategy

Starlab®

A Lion

Starlab®

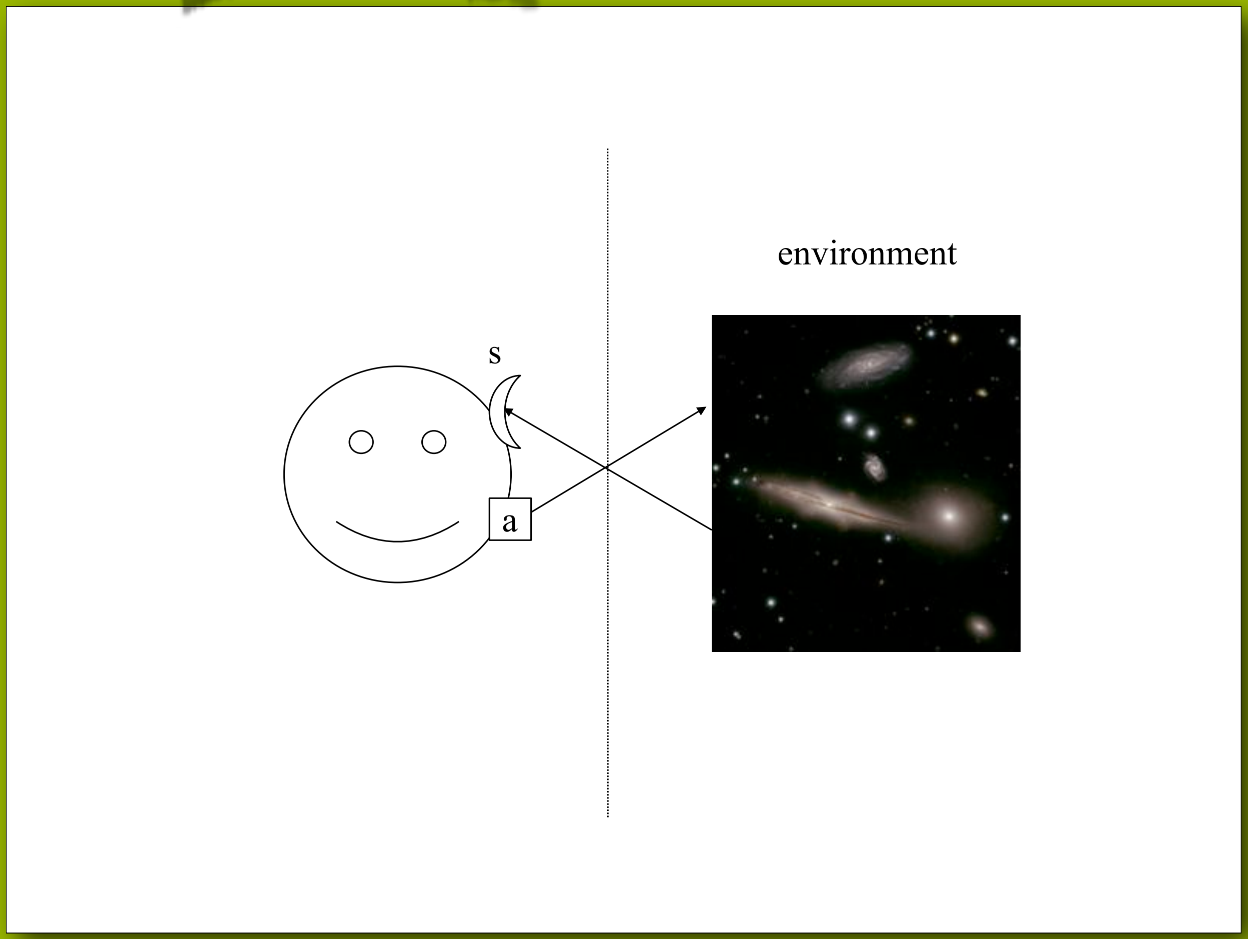
Premises

1. Human experience is generated in the brain
2. All the brain has access to is information

This would clearly apply to the CPU in a robot. Robot body parts are part of the environment to model (self-model)

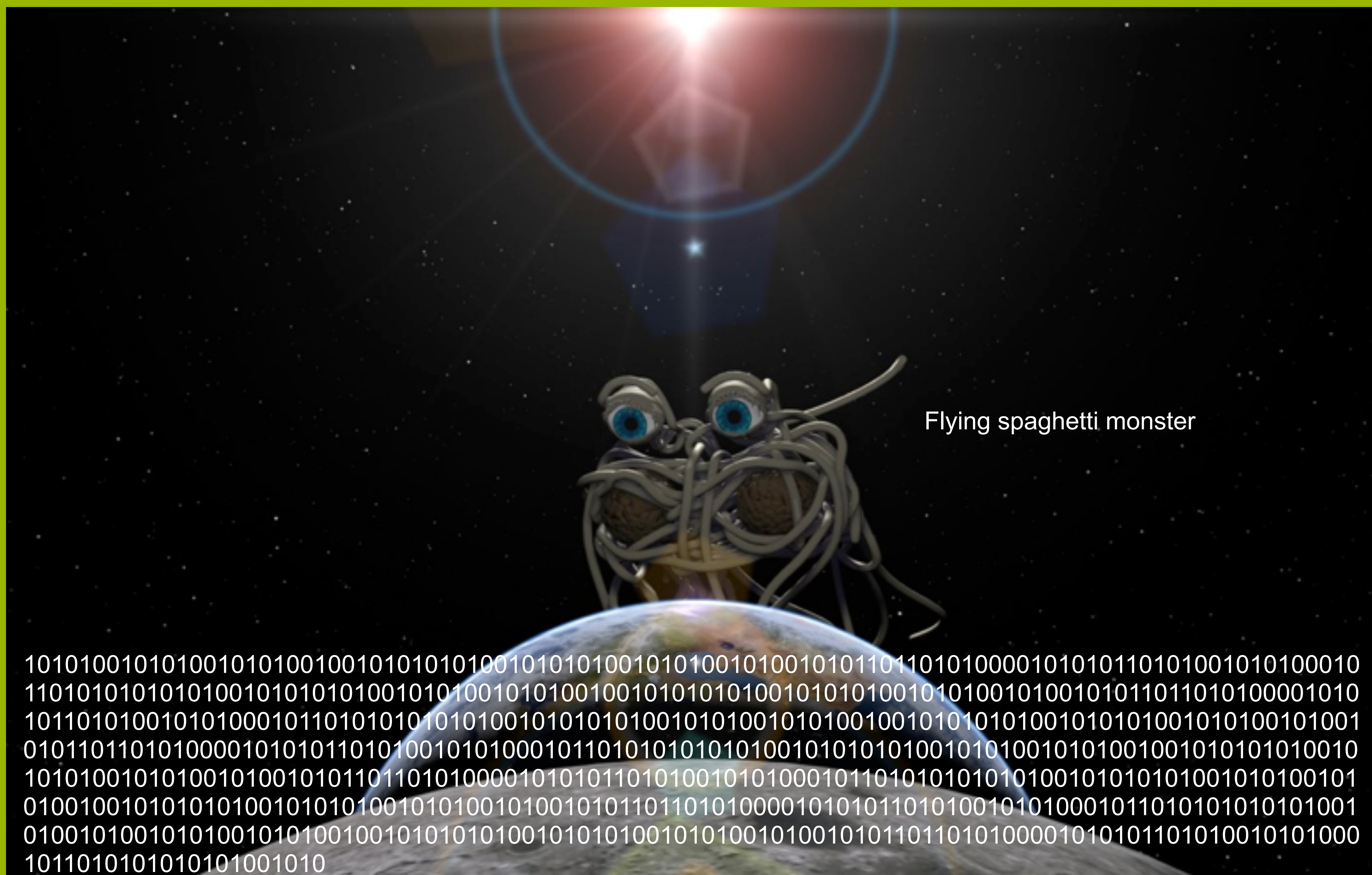
How extreme is this view?

- Information is *the* fundamental concept in physics today. From atoms, to quanta, to bits?
- States and dynamics: information and computation
- "It from bit" (John Wheeler). Information as the ultimate building block. As far as we know, this is *it*.
- The Presence *gedanken*: ideal VR experiment shows all that is needed is bits



Starlab®

Let there be bits!



X

Starlab®

Models = Algorithms

- A model is a data compressor
- A model is a data predictor

- A model uses physical resources: infrastructure, energy, time
- Recursive computable functions a paradigm of models: algorithms. Turing machines.

- Warning 1: a computer is needed!
- Warning 2: time is need!

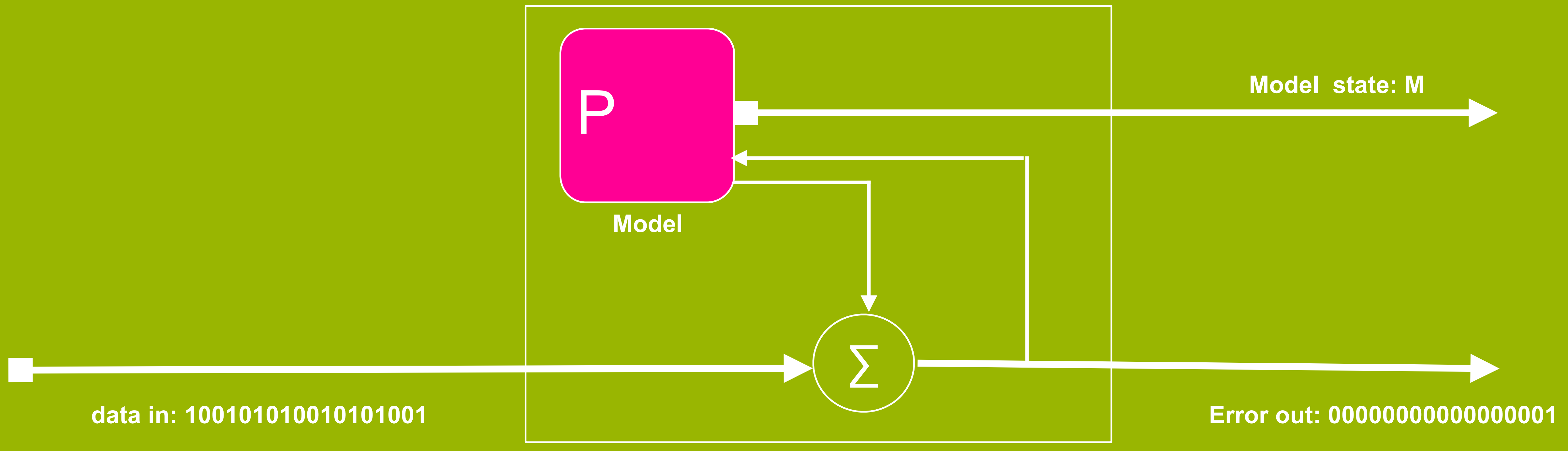
Examples of models

- Conservation of mass
- Conservation of charge
- Tigers
- People
- Newton's laws
- Quantum physics
- Time
- Your body
- You

High algorithmic mutual information with Environment

Starlab®

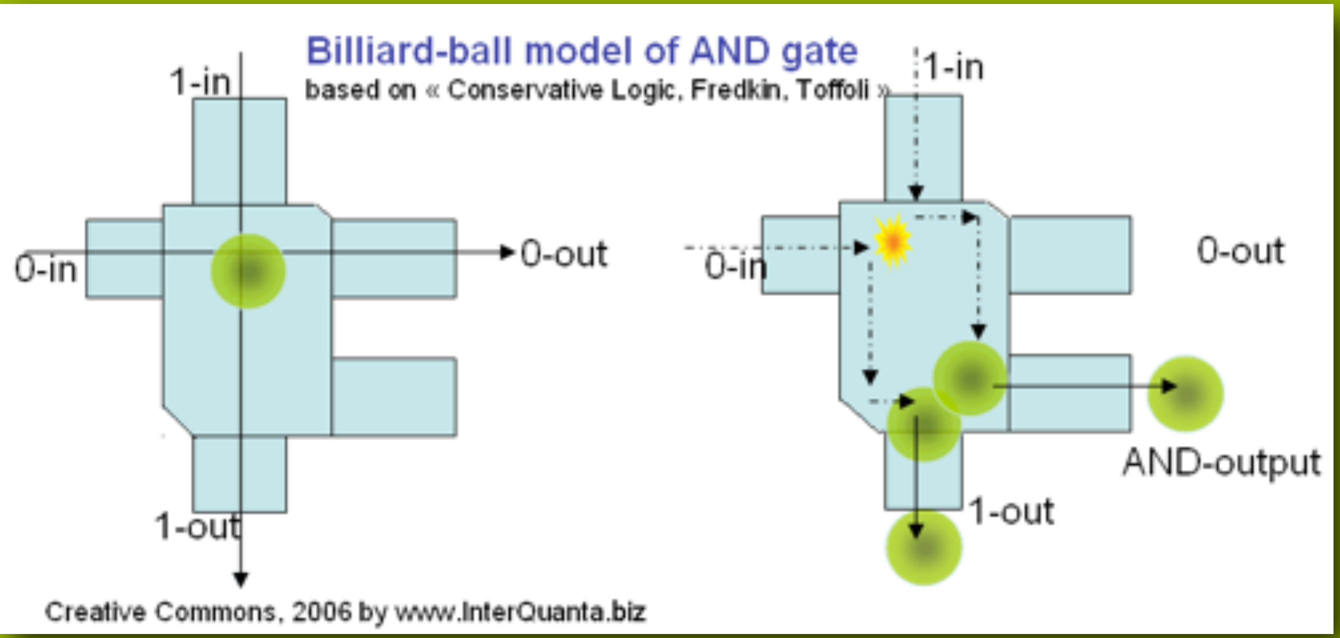
Models at work: predictive coding



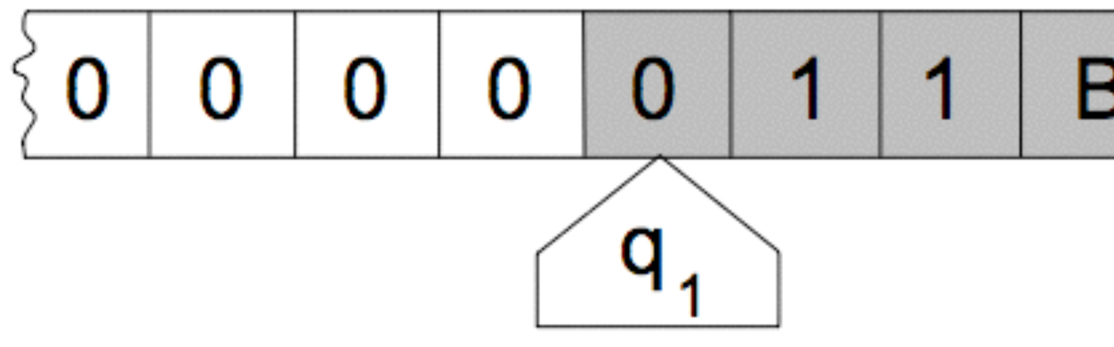
Starlab®

Computation = dynamics

- Computation= a set of states+ dynamics
- Classical, quantum, Turing ... all based on these



Newtonian dynamics



Turing machine

Starlab®

Inference machines (2008)

Physical limits of inference

David H. Wolpert

MS 269-1, NASA Ames Research Center, Moffett Field, CA 94035, USA

Abstract

I show that physical devices that perform observation, prediction, or recollection share an underlying mathematical structure. I call devices with that structure "inference devices". I present a set of existence and impossibility results concerning inference devices. These results hold independent of the precise physical laws governing our universe. In a limited sense, the impossibility results establish that Laplace was wrong to claim that even in a classical, non-chaotic universe the future can be unerringly predicted, given sufficient knowledge of the present. Alternatively, these impossibility results can be viewed as a non-quantum mechanical "uncertainty principle". Next I explore the close connections between the mathematics of inference devices and of Turing Machines. In particular, the impossibility results for inference devices are similar to the Halting theorem for TM's. Furthermore, one can define an analog of Universal TM's (UTM's) for inference devices. I call those analogs "strong inference devices". I use strong inference devices to define the "inference complexity" of an inference task, which is the analog of the Kolmogorov complexity of computing a string. However no universe can contain more than one strong inference device. So whereas the Kolmogorov complexity of a string is arbitrary up to specification of the UTM, there is no such arbitrariness in the inference complexity of an inference task. I end by discussing the philosophical implications of these results, e.g., for whether the universe "is" a computer.

Key words: Turing machine, automata, observation, prediction, multiverse, Kolmogorov complexity
PACS: 03.65.Ta, 89.20.Ff, 02.70.-c, 07.05.Tp, 89.70.Eg, 01.70.+w

Starlab®

The ultimate laptop (Lloyd 2000)

Figure 1 The ultimate laptop. The 'ultimate laptop' is a computer with a mass of 1 kg and a volume of 1 l, operating at the fundamental limits of speed and memory capacity fixed by physics. The ultimate laptop performs $2mc^2/\pi\hbar = 5.4258 \times 10^{50}$ logical operations per second on $\sim 10^{31}$ bits. Although its computational machinery is in fact in a highly specified physical state with zero entropy, while it performs a computation that uses all its resources of energy and memory space it appears to an outside observer to be in a thermal state at $\sim 10^9$ degrees Kelvin. The ultimate laptop looks like a small piece of the Big Bang.

Ultimate physical limits to computation

Seth Lloyd

d'Arbeloff Laboratory for Information Systems and Technology, MIT Department of Mechanical Engineering, Massachusetts Institute of Technology 3-160, Cambridge, Massachusetts 02139, USA (slloyd@mit.edu)

Is the universe a computer?

- The universe is described by a physical state (a number in some representation)
- Obeys some dynamics
- Has some computational power ... it can be estimated!

- Any physical system can be emulated by a (quantum) computer

Starlab®

How powerful is it?

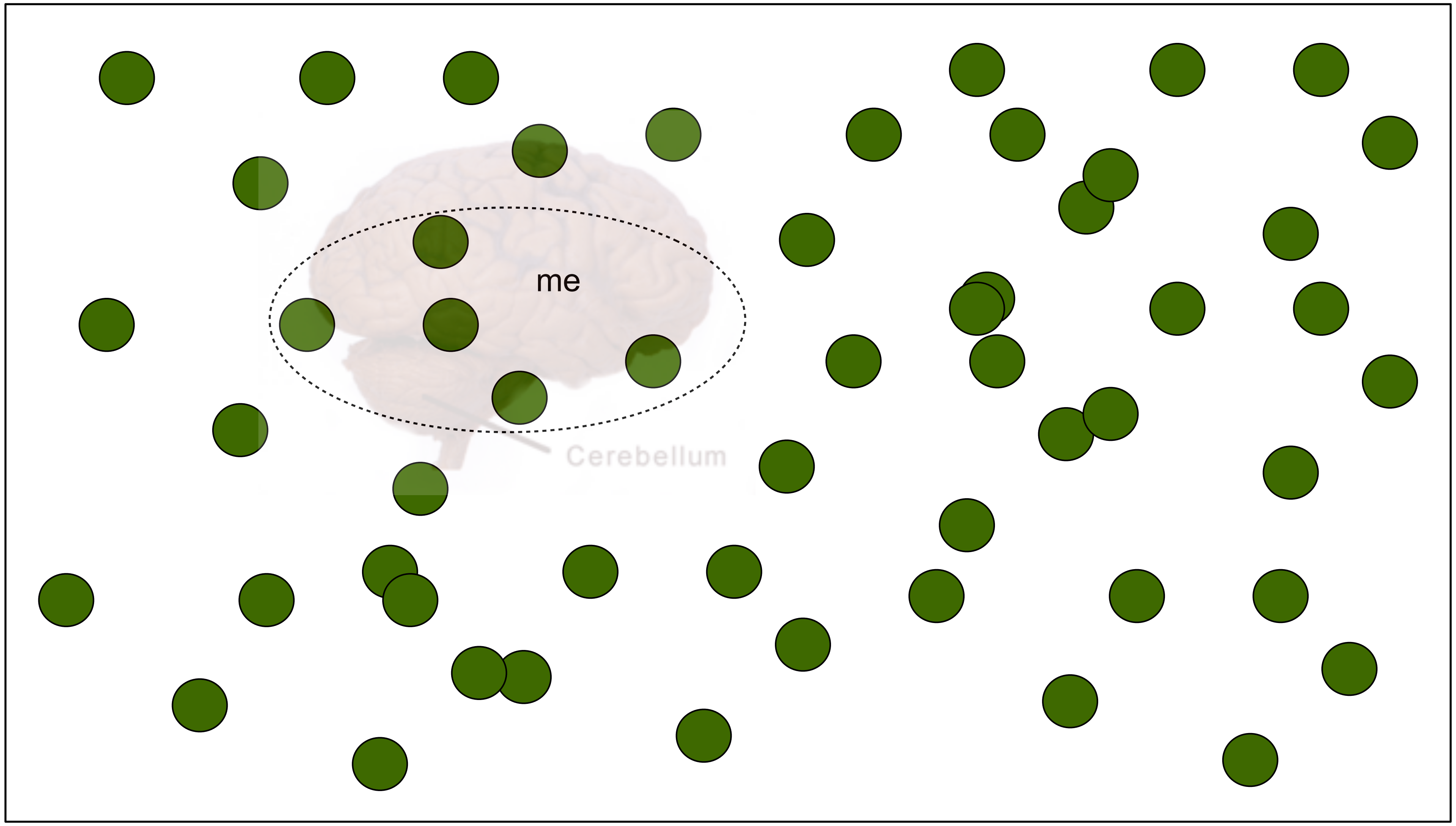
Computational capacity of the universe

Seth Lloyd
d'Arbeloff Laboratory for Information Systems and Technology
MIT Department of Mechanical Engineering
MIT 3-160, Cambridge, Mass. 02139
slloyd@mit.edu

Merely by existing, all physical systems register information. And by evolving dynamically in time, they transform and process that information. The laws of physics determine the amount of information that a physical system can register (number of bits) and the number of elementary logic operations that a system can perform (number of ops). The universe is a physical system. This paper quantifies the amount of information that the universe can register and the number of elementary operations that it can have performed over its history. The universe can have performed no more than 10^{120} ops on 10^{90} bits.

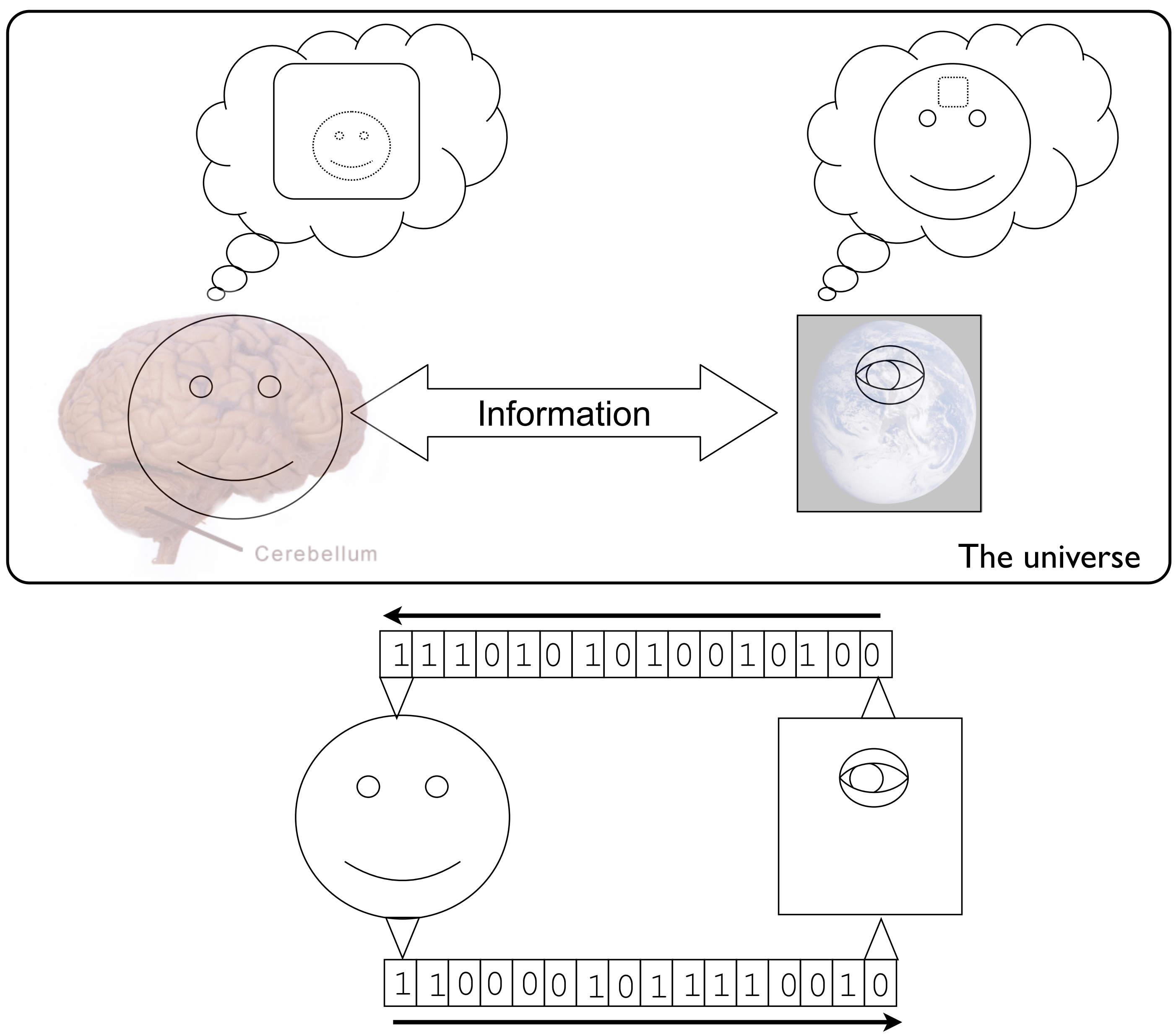
But where is the boundary?

The Universe

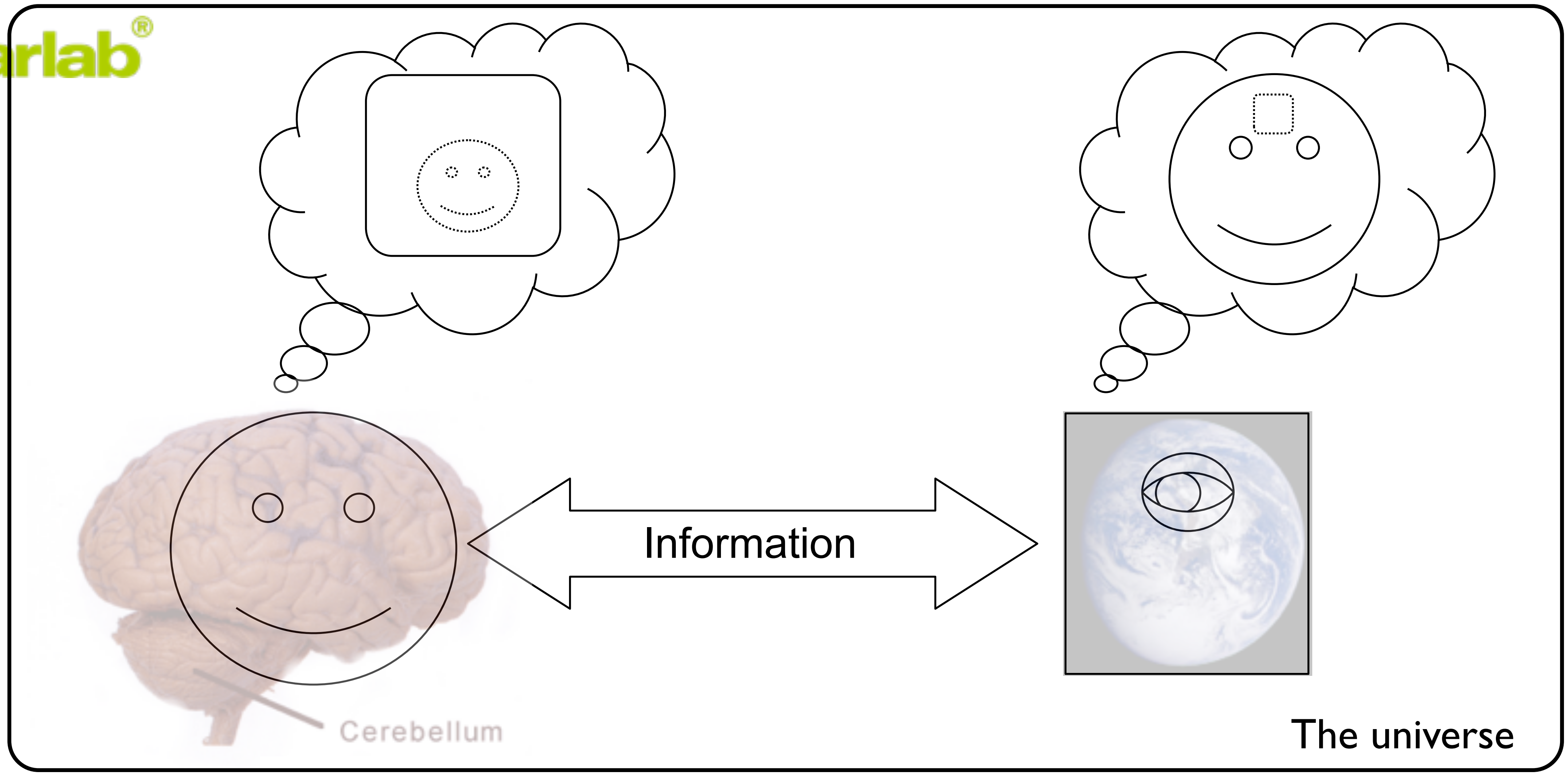


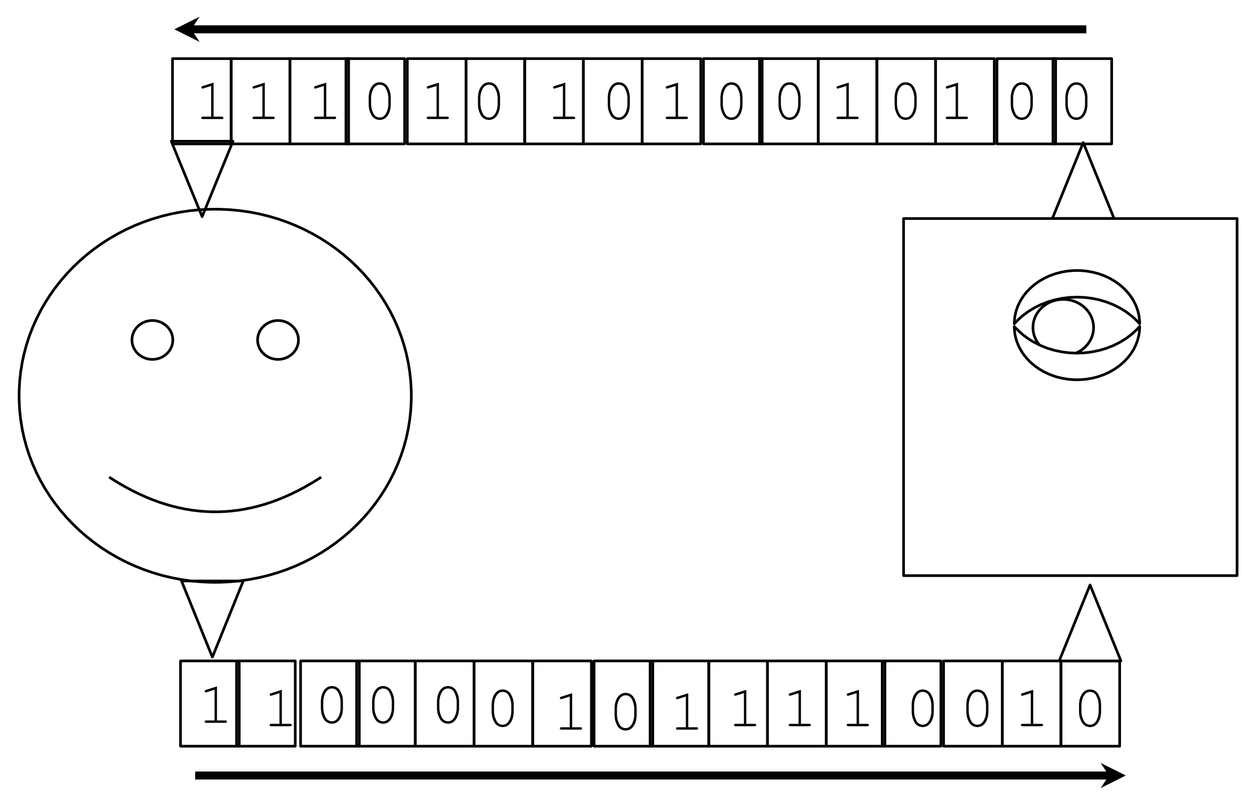
Starlab®

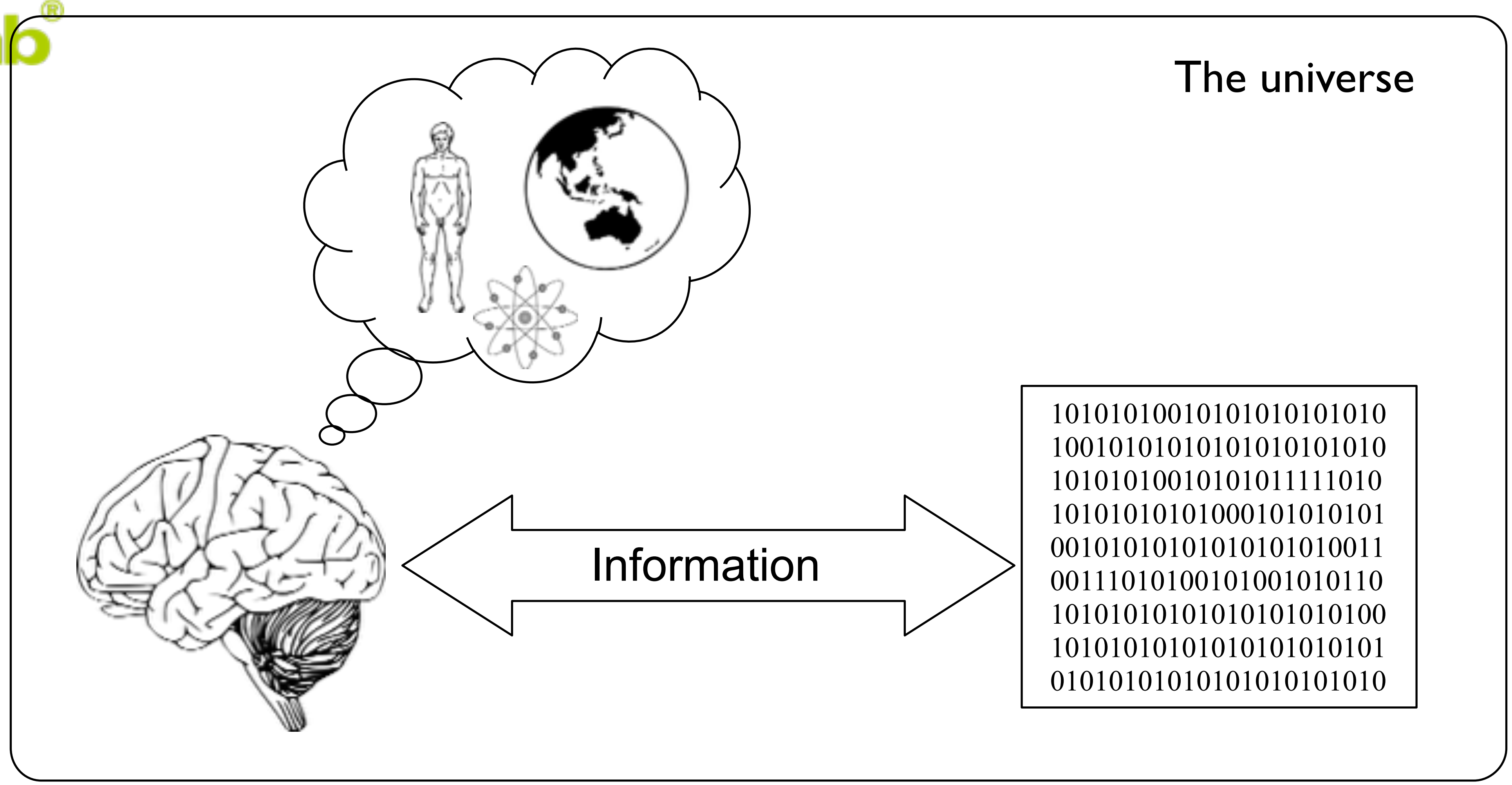
How does it happen?

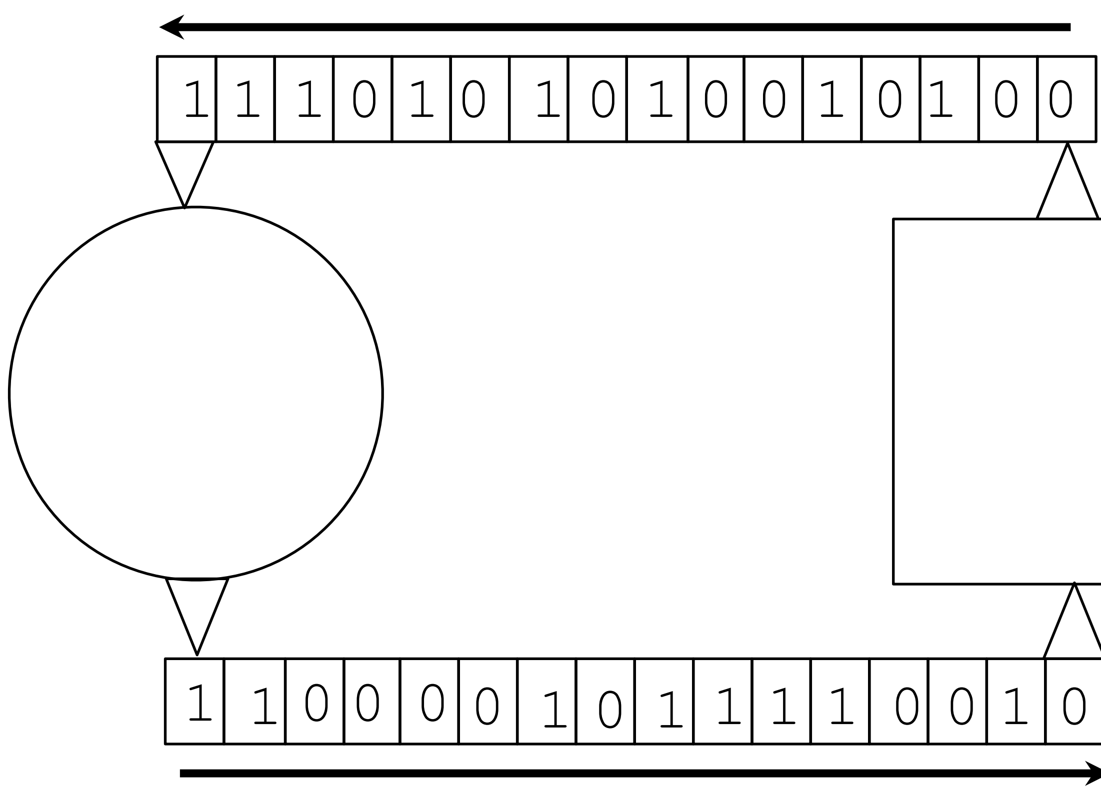


Starlab®

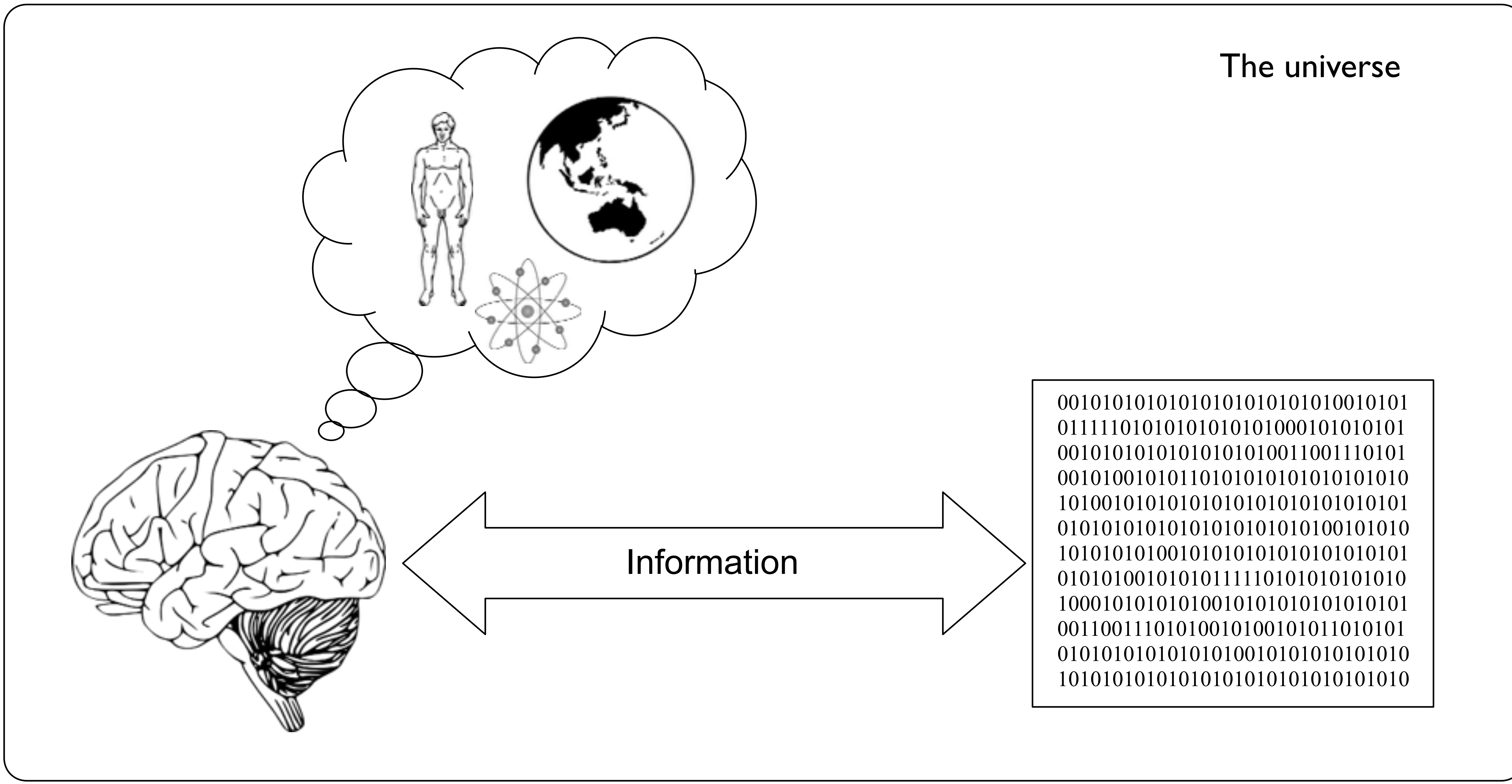

Information
Cerebellum
The universe


1 1 1 0 1 0 1 0 1 0 0 1 0 1 0 0
1 1 0 0 0 0 1 0 1 1 1 1 0 0 1 0

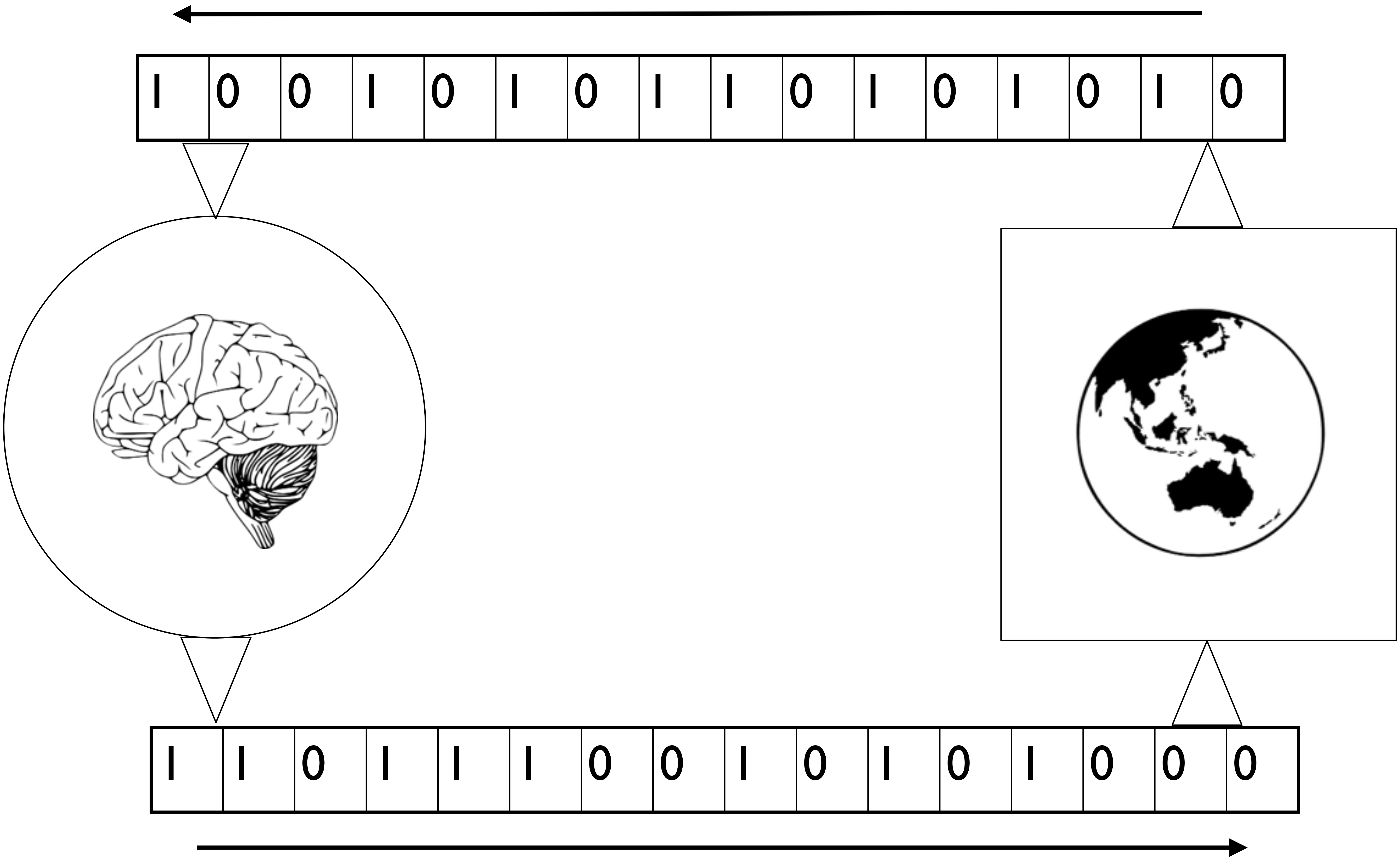

The universe
Information
10101010010101010101010
10010101010101010101010
10101010010101011111010
10101010101000101010101
00101010101010101010011
00111010100101001010110
10101010101010101010100
10101010101010101010101
01010101010101010101010

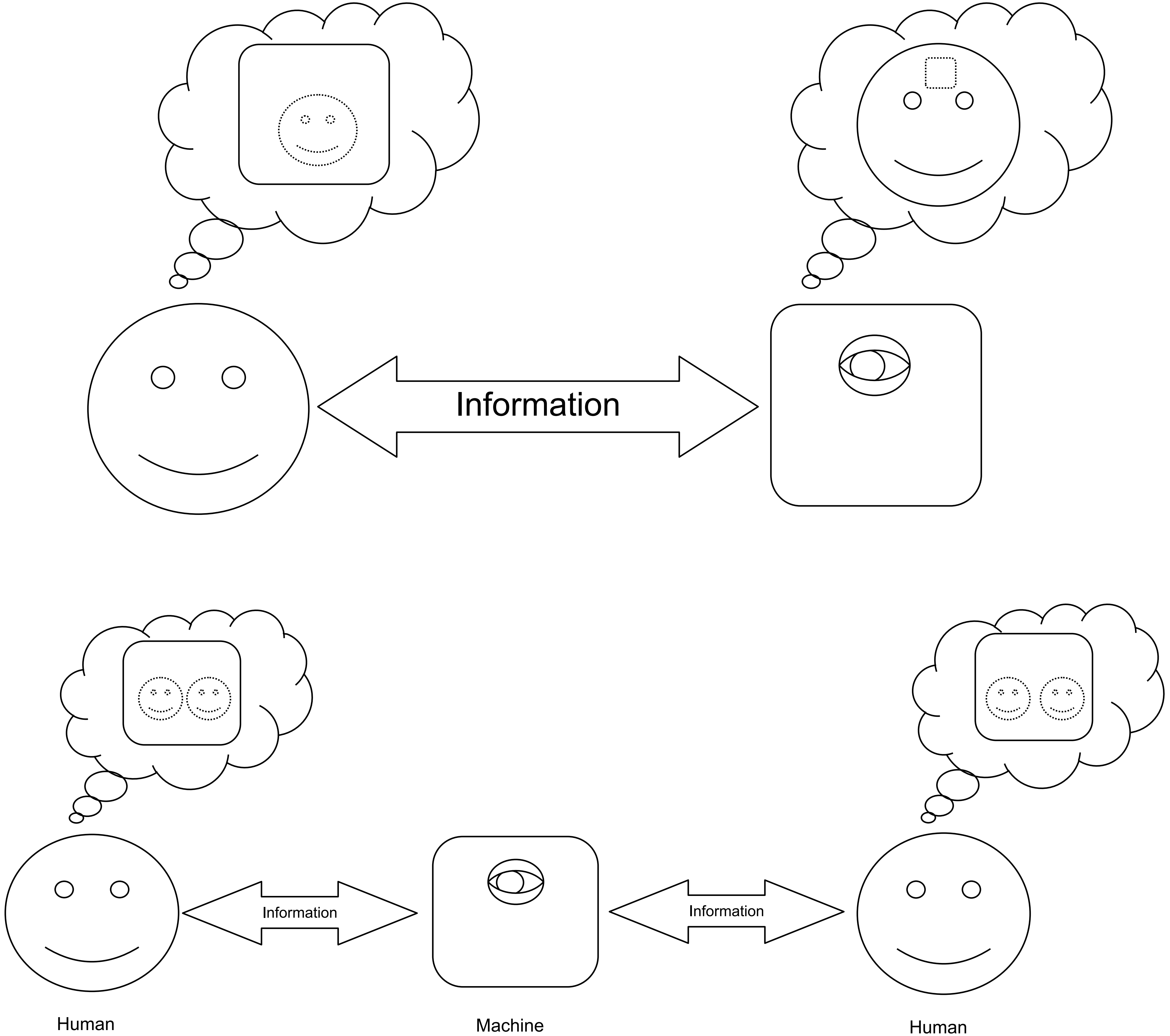

1 1 1 0 1 0 1 0 1 0 0 1 0 1 0 0
1 1 0 0 0 0 1 0 1 1 1 1 0 0 1 0

Starlab®


The universe
Information
00101010101010101010101010010101
01111101010101010101000101010101
00101010101010101010011001110101
00101001010110101010101010101010
10100101010101010101010101010101
01010101010101010101010100101010
10101010100101010101010101010101
01010100101010111110101010101010
10001010101010010101010101010101
00110011101010010100101011010101
01010101010101010010101010101010
10101010101010101010101010101010

Starlab®





Information
Information
Information
Human
Machine
Human

Simplicity

Starlab®

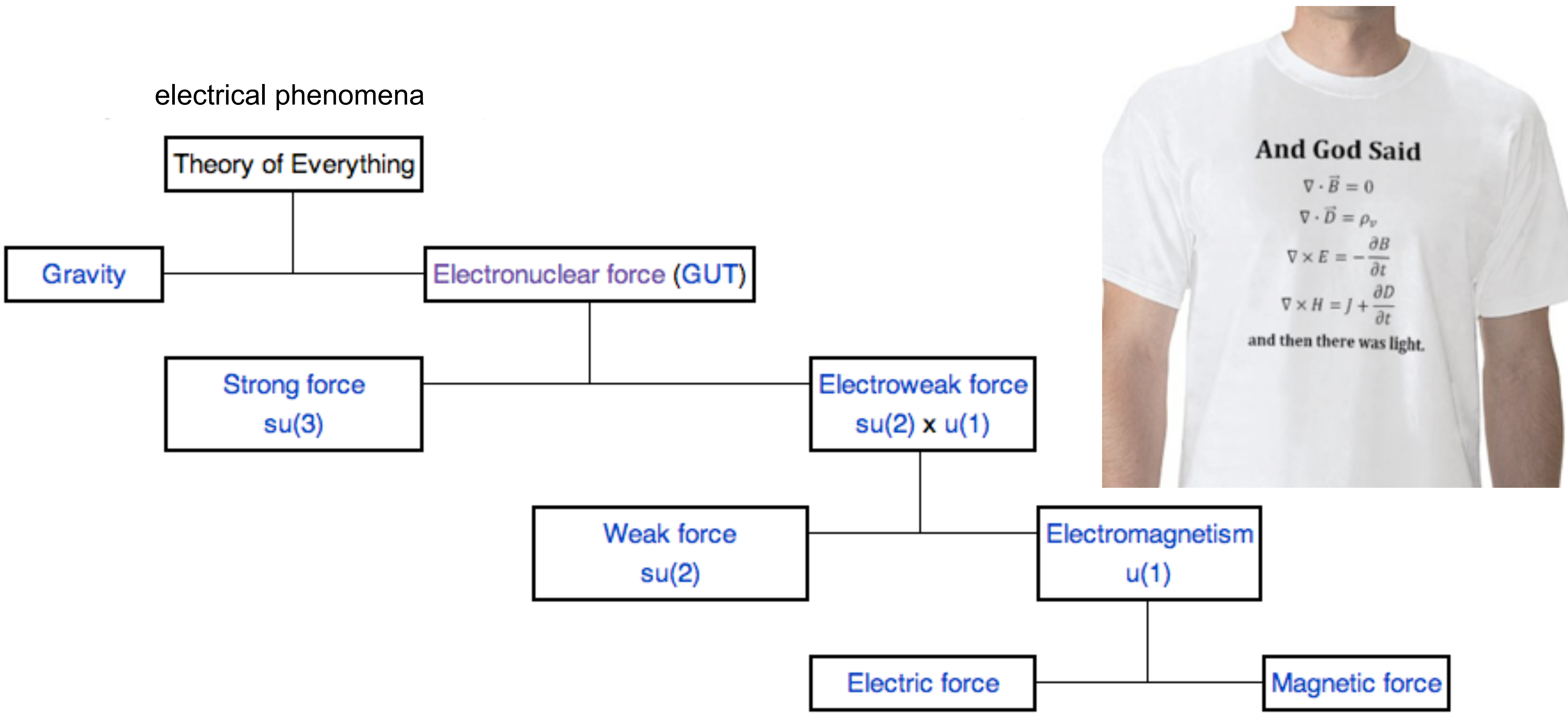
Simplicity in Science

- Science is the search for simplicity in the available (i.e., partial) information we have access to.
- Going from data to models = compression
- Science's goal is to develop models
  - Models must account for all data in an economic way
  - Models must predict future data well

- Simple models appear to be superior. Why is this?
- Example: TOE

Compression in Physics

electrical phenomena



Starlab®

Algorithmic Complexity

- Compression and therefore simplicity were first successfully formalized by the notion of algorithmic complexity or Kolmogorov complexity (also known as `algorithmic information', `algorithmic entropy', `Kolmogorov-Chaitin complexity', `descriptional complexity', `shortest program length' and `algorithmic randomness'.)
- Co-discovered during the second half of the 20th century by Solomonoff, Kolmogorov and Chaitin
- Provides a well-established albeit formal cornerstone to address the question of compression in brains---both natural or artificial.
- We recall its definition: loosely, the Kolmogorov complexity of a data set is the length of the shortest program capable of generating it.

Starlab®

More precisely

generating it. More precisely (see e.g., [14] or [15]), let U be a universal computer (a Turing machine), and let p be a program. Then the Kolmogorov or algorithmic complexity of a string x with respect to U is defined by

$$K_U(x) = \min_{p:\, U(p)=x} l(p), \qquad (1)$$

the minimum length over all programs that print the string x and then halt. The restriction to programs that halt is important, for no program is the concatenation of other programs. An important fact is that this is a meaningful definition: although the precise length of the minimizing program depends on the programming language used, it does so only up to a constant. That is, if U is a universal computer, then for any computer A we can easily show that $K_U(x) < K_A(x) + c$. The constant c is the length of the program for U to emulate A.

Gödel's incompleteness theorem implies we cannot compute in general the KC of an arbitrary string. As is discussed in [16], Gödel's theorem is

Examples

As a first example of this concept, consider the sequence

```
1212121212121212121212121212121212121212
1212121212121212121212121212121212121212
```

It is easy to see that its Kolmogorov Complexity is rather small. Here is a simple algorithm to describe it:

“`Repeat 12 forty times`”

```
4811174502 8410270193 8521105559 6446229489 5493038196
4428810975 6659334461 2847564823 3786783165 2712019091
4564856692 3460348610 4543266482 1339360726 0249141273
7245870066 0631558817 4881520920 9628292540 9171536436
7892590360 0113305305 4882046652 1384146951 9415116094
3305727036 5759591953 0921861173 8193261179 3105118548.
```

digits 151 to 450 of pi

A small problem

- There is no algorithm to compute K
- This is due to Godel's incompleteness theorem, or equivalently, the halting problem (Turing)
- Cannot test programs ... no assurance that they will ever stop
- However, in practice, if there is a finite time limit, we can compromise

Starlab®

A related approach

- Algorithmic probability
- Suppose you are given a string x. What is the probability that a monkey typing on a Turing machine would have generated it?

A further connection to simplicity using the concept of Kolmogorov or algorithmic complexity was developed by Solomonoff [15] with an emphasis on prediction. The fundamental concept is the algorithmic or universal (un-normalized) probability $P_U(x)$ of a string x. This is the probability that a given string x could be generated by a random program. It is given by

$$P_U(x) = \sum_{p:\, U(p)=x} 2^{-l(p)} \tag{2}$$

A very important result connecting this probability to complexity theory is that

$$P_U(x) \approx 2^{-K_U(x)}, \tag{3}$$

that is, the probability of a given string to be produced by a random program is dominated by its Kolmogorov complexity. Here the earlier remark that

Starlab®

MDL

- Minimum description length is a ML algorithm for statistical inference.
- The sum of program length plus “error out” length is minimized.
- Tradeoff between program length and good of fit
- Nicely related to Bayesian and KC approaches

- Use for inference relies on onnection mediated by a “prior”. This is the universal prior of Solomonoff

Bayes and Occam's razor

$$p(a,b) = p(a|b)\,p(b) = p(b|a)\,p(a)$$

The relation of Ockham's razor to Bayesian theory is discussed for example in [19], [18] and [4]. Given two models, b_1 and b_2, their relative probability given some data a is

$$\frac{p(b_1|a)}{p(b_2|a)} = \frac{p(b_1)\,p(a|b_1)}{p(b_2)\,p(a|b_2)}. \tag{9}$$

Without access to any prior information,

$$\frac{p(b_1|a)}{p(b_2|a)} \sim \frac{p(a|b_1)}{p(a|b_2)}. \tag{10}$$

Starlab®

Jaynes and MEP

- Probability as an extension of logic
- Simple models represent in a less biased way our knowledge of the exterior world
- Principle of indifference (Laplace). If we have no prior information on a set of scenarios, assign equal probabilities
- Problem: there is no canonic way to split the space of scenarios.

Starlab®

MDL, Bayes and complexity

Minimum Description Length Induction, Bayesianism, and Kolmogorov Complexity

Paul M. B. Vitányi and Ming Li

***Abstract*—The relationship between the Bayesian approach and the minimum description length approach is established. We sharpen and clarify the general modeling principles minimum description length (MDL) and minimum message length (MML), abstracted as the ideal MDL principle and defined from Bayes's rule by means of Kolmogorov complexity. The basic condition under which the ideal principle should be applied is encapsulated as the fundamental inequality, which in broad terms states that the principle is valid when the data are random, relative to every contemplated hypothesis and also these hypotheses are random relative to the (universal) prior. The ideal principle states that the prior probability associated with the hypothesis should be given by the algorithmic universal probability, and the sum of the log universal probability of the model plus the log of the probability of the data given the model should be minimized. If we restrict the model class to finite sets then application of the ideal principle turns into Kolmogorov's minimal sufficient statistic. In general, we show that data compression is almost always the best strategy, both in model selection and prediction.**

Starlab®

Inference - No free lunch!

IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, VOL. 1, NO. 1, APRIL 1997 67

No Free Lunch Theorems for Optimization

David H. Wolpert and William G. Macready

***Abstract*—A framework is developed to explore the connection between effective optimization algorithms and the problems they are solving. A number of "no free lunch" (NFL) theorems are presented which establish that for any algorithm, any elevated performance over one class of problems is offset by performance over another class. These theorems result in a geometric interpretation of what it means for an algorithm to be well suited to an optimization problem. Applications of the NFL theorems to information-theoretic aspects of optimization and benchmark measures of performance are also presented. Other issues addressed include time-varying optimization problems and *a priori* "head-to-head" minimax distinctions between optimization algorithms, distinctions that result despite the NFL theorems' enforcing of a type of uniformity over all algorithms.**

***Index Terms*— Evolutionary algorithms, information theory, optimization.**

Starlab®

Why is simplicity important?

Starlab®

A1: Knowledge representation

- Suppose Axioms **A**, **B** and **C** "explain" (decompress to) the facts.
- You can also add **D** if it does not conflict with the facts. But adding it is a disservice to your representation of knowledge.

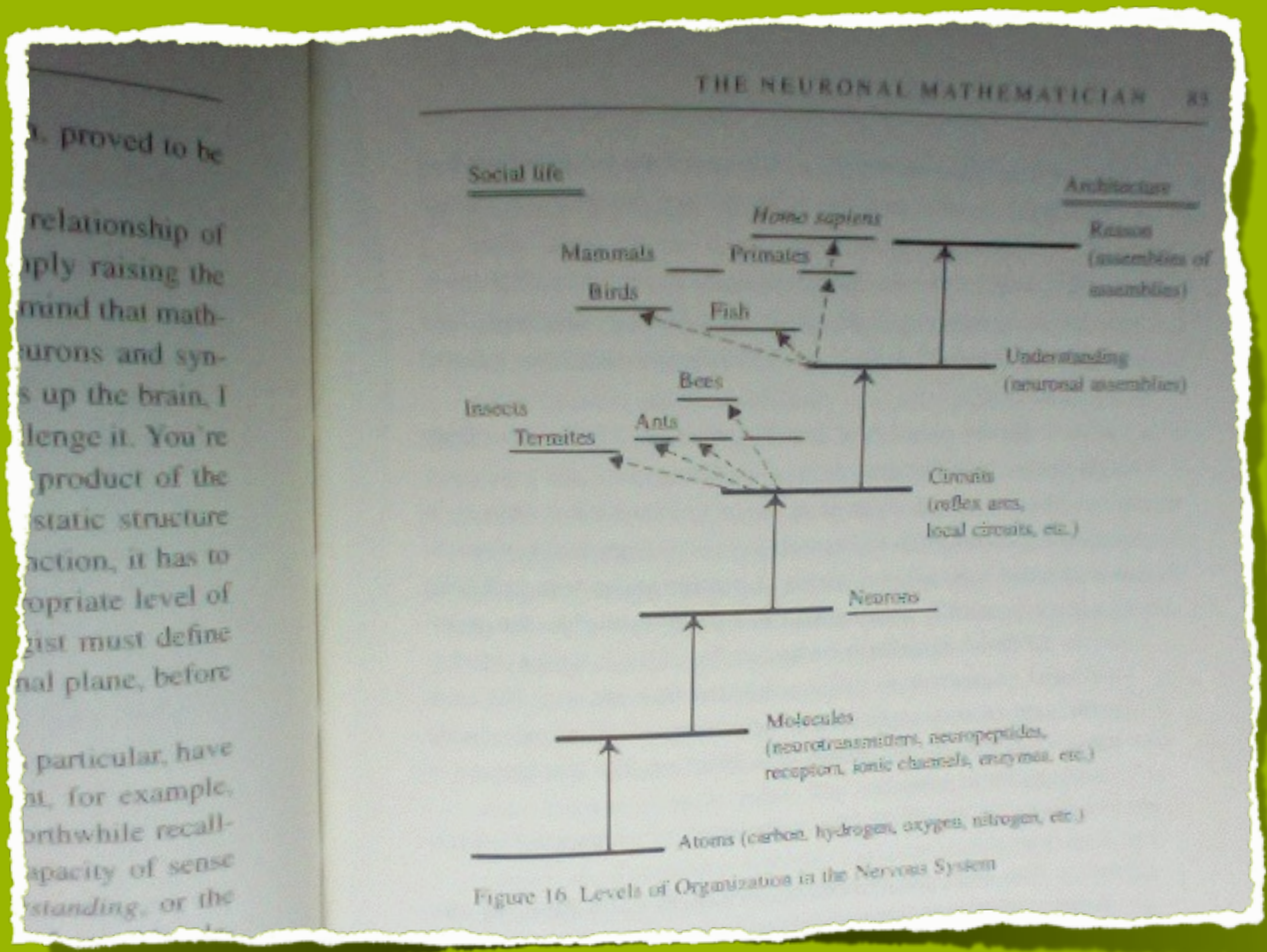
Starlab®

A2: Evolution

- Simplicity and natural selection can guide us
- Recalling: using a model data is compressed using finite resources. Makes sense is such models exist
- Acting: experimenting observing more efficient if there is a baseline model to start with; simple models easier to run and use for decision making. Homeostasis/pain
- Predicting: if models do exist, then prediction is possible, and this clearly helps. If the universe is simple ... everything follows.
- If the universe is not simple ...
  - Evolution leads to layered simplifying processors: start simple
  - Simple models represent what we know in the most economical way, easier to use for inference and deduction
  - The simpler the better: also for debugging!

Starlab®

Hierarchies



The brain contains the uppermost modeling hierarchies. Short time scales learning.

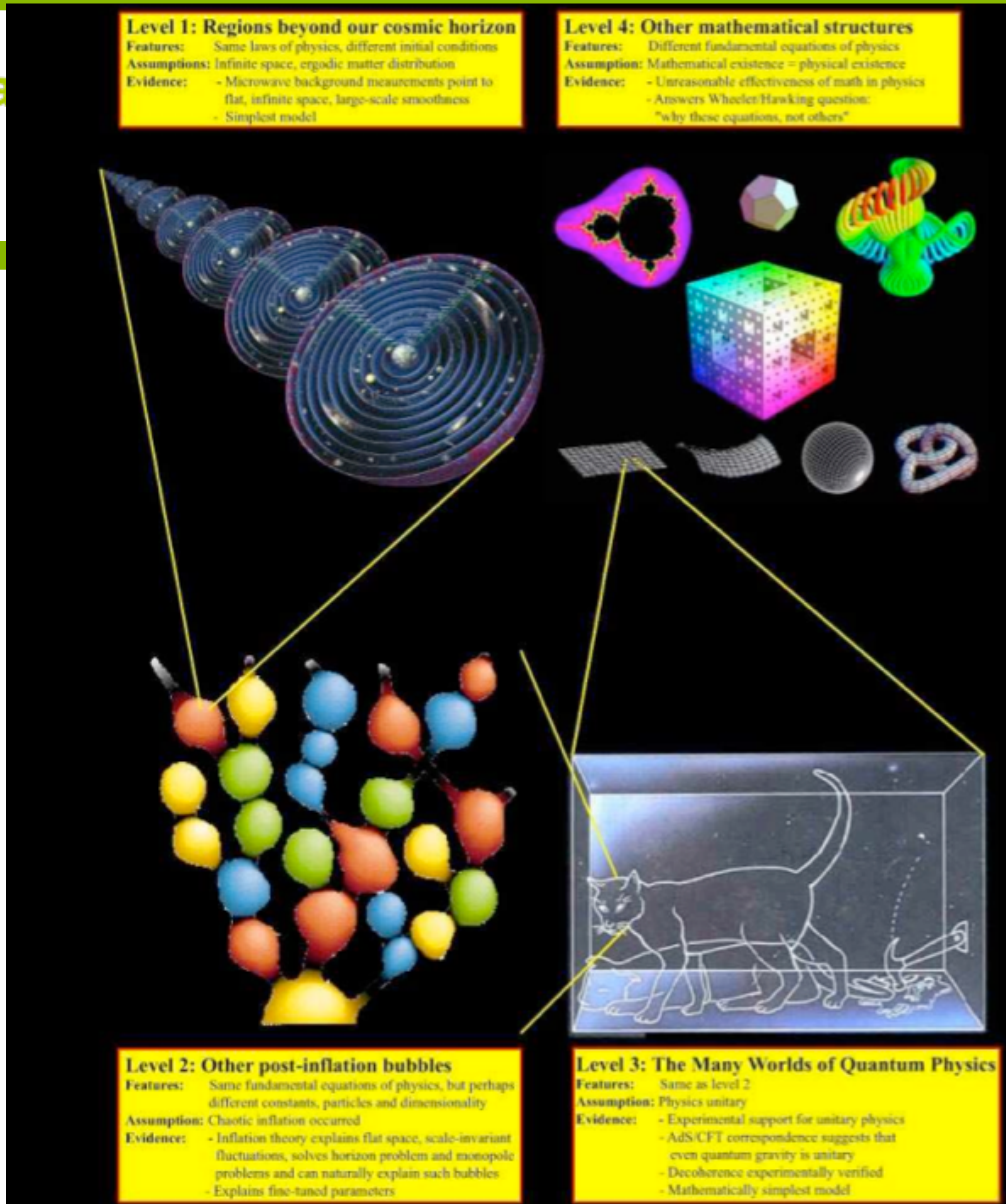
The body, the organism level, is at a lower hierarchy, "learning" taking place at longer time scales -- eons.

Starlab®

A3: Inferotropic Principle

- Modeling simply may be the best practical strategy. But why does it work so well? Is the universe simple?
  - The universe is simple because it arises from random programs
  - *Inferotropism*
    - The universe is simple because inference machines can only exist in simple universes.
    - There are many universes, or we only collect a subset of the available data ... the one which can actually sustain simplicity and inference machines

Sta



Universe or Multiverse? Which?

The Multiverse Hierarchy

Max Tegmark

Dept. of Physics & MIT Kavli Institute, Massachusetts Institute of Technology, Cambridge, MA 02139

ted: February 20 2005, published in *"Universe or Multiverse?"*, B. Carr ed., Cambridge University Press, 2007))

X

Is this the answer?

The relevance of simplicity is also at the core of the question of why mathematics and simple theories turn out to be the right ones in physics and science—a mystery that has perplexed many thinkers, including Einstein and Feynman. A possible answer is that inference machines will by their nature always seek and find some simplicity, even in random data. Let us recall here that any desired sequence can be found in a truly random number. An inference machine exposed to a lucky portion of the stream will deduce all sorts of things which may just be a passing mirage. This is the complexity, or rather simplicity, version of the Anthropic principle (see e.g., [23]). We could also say that without simplicity there would not be infererence machines.

- Or is god a monkey typing away on Turing machine to ensure that algorithmic probability rules?

Starlab®

Applications in neuroscience

Starlab®

Simplicity in brains

- If reality is just a model in the brain,
  - and if simple models are better than complex models,
  - then *reality* = simplicity
- But is this really true? Can we test it?

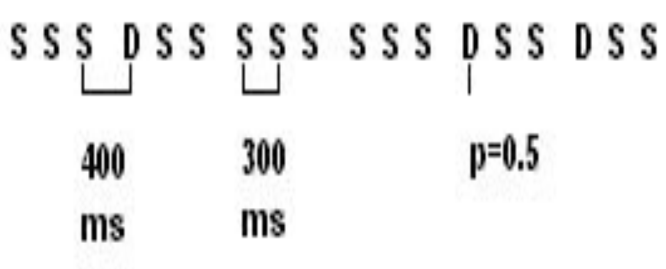
Starlab®

ERPs and MMN

- Event-Related Potentials, ERPs, are a useful non-invasive window for the study of fast (ms) cerebral processes.
- ERPs involve searching for event-locked regularities of brain dynamics, averaging multiple time intervals that share the same experimental conditions.
- Mismatch negativity (MMN) is an involuntary auditory ERP which peaks at 100–200 ms when there is a violation (deviant tone) of a regular pattern (standard tone sequence).
- The MMN mechanism appears to correspond to a primitive intelligence, as the wave produced with the violation of regularities—even those of an abstract nature

Standard and Deviant

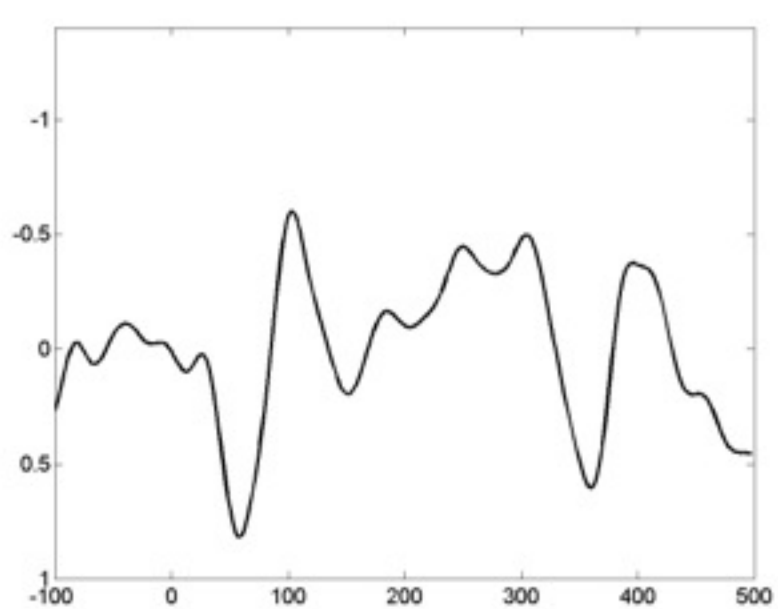
- Binaural stimulation
- EEG recorded using 30 electrodes with a nose reference plus 2 EOG



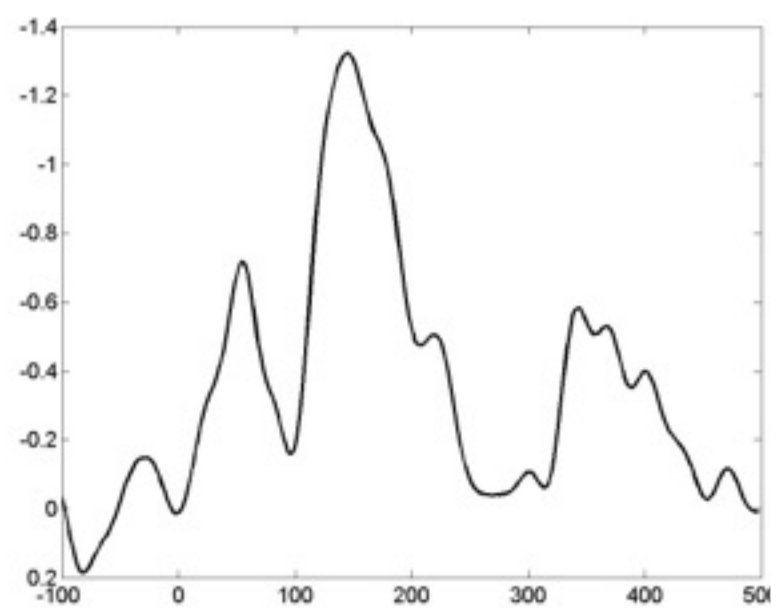
700 Hz, 85 dB SPL

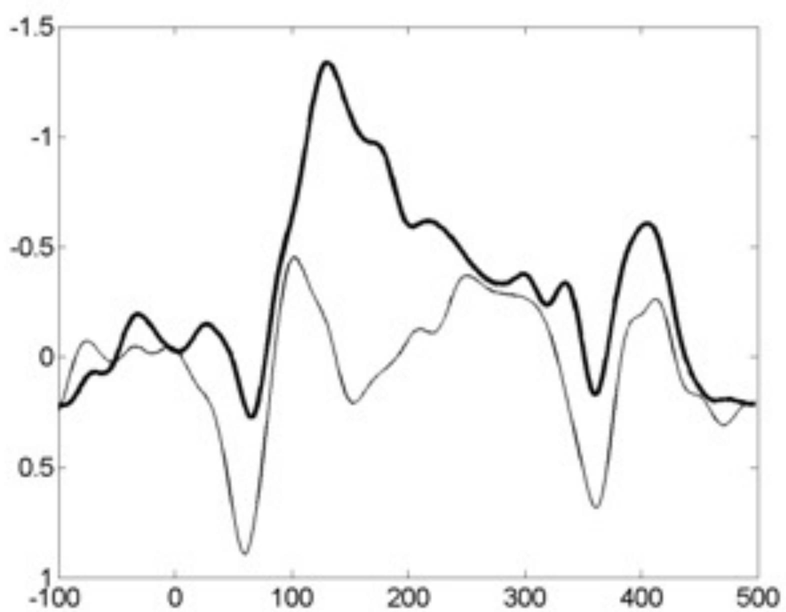
D = deviant, 75 ms
S = standard, 25 ms

Standard



Deviant





ABAB ABAB ABAB BABA

- This sequence does not generate MMN
  - ABAB ABAB ABAB **B**ABA **A**BAB ABAB ABAB **B**ABA

- The MMN does not activate. Or is there a generic rule used? “Alternate within a block”.

$$S_{n+1} = NOT(S_n)$$

$$S1 = R_0(R_3(AB) + R_n(-) + R_3(BA) + R_n(-))$$

ABABABABABABBABA

- This sequence does generate MMN:
  - ABABABABABAB**B**ABA**A**BABABABABAB**B**ABA
- It is the same sequence as before, but spaces are removed. *Alternation* rule is violated.

$$S_{n+1} = NOT(S_n)$$

Test model building using MMN

- W can try to use complexity measures to explore model building in MMN - future work.
- Is MMN at the bottom of a hierarchy?

Z1= ABABAB - - ABABAB - - ABABAB ...
Z2= AABABA - - AABABA - - AABABA ...
Z3= BAABAB - - BAABAB - - BAABAB ...

$$Z1 = R_0(R_3(AB) + R_n(-))$$
$$Z3 = R_0(R_1(BA) + R_2(AB) + R_n(-))$$

Music and MMN

- Music may be a phenomenon associated to model building: looking for simplicity in data.

Starlab®

Application to Presence

Different levels of Presence

RAVE: pi+psi

10 Mapping our brains to computers (the singularity)
9 Jacking in (invasive interaction)
8 Non-invasive Brain 2 Machine + Machine 2 Brain interaction
7 Immersion (HMD/CAVE + haptics + ...) (also MR/AR) using natural senses

_____________________*FUZZY Raving DIVIDE*_____________________

6 Disneyworld; 2nd Life;
5 Cinema/IMAX; telepresence
4 Theater
3 Books, mobile phones, chatting
2 Storytelling
1 Imagery

Starlab®

Hierarchies

- From transducers to social phenomena
- MMN at the bottom, for example.
- In Presence research, we see evidence for such hierarchies: e.g., sensorimotor contingencies=low level model consistency.
- Two recently proposed Presence dimensions (Slater)
  - PI: Place Illusion (M. Slater)
  - Psi: Plausibility

- Both are part of the same picture: model building at different levels of the hierarchy.

Starlab®

Modeling, PI and Psi

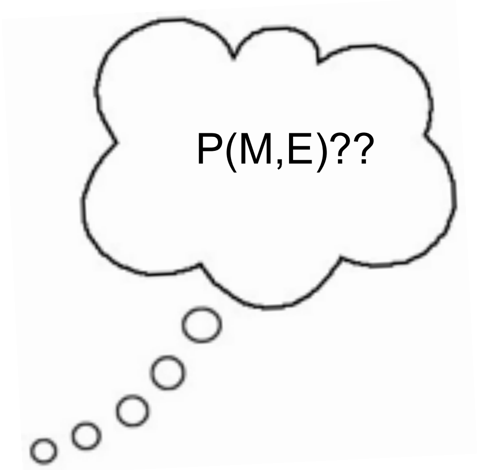
- *PI*: the strong illusion of being in a place in spite of the sure knowledge that you are not there: low level modeling
- *Psi:* the illusion that what is apparently happening is real ly happening (even though you know for sure that it is not).
- Thus, Psi represents an escalation in the creation of illusions higher up (but not to the top) in the modeling hierarchy.
- From the point of experience design, we could also add that what is intended to appear to be happening is the actual perceived illusion.
- The ultimate level in this context may be called “Susi” -- “suckership”: believing all the way, “being sucked in”

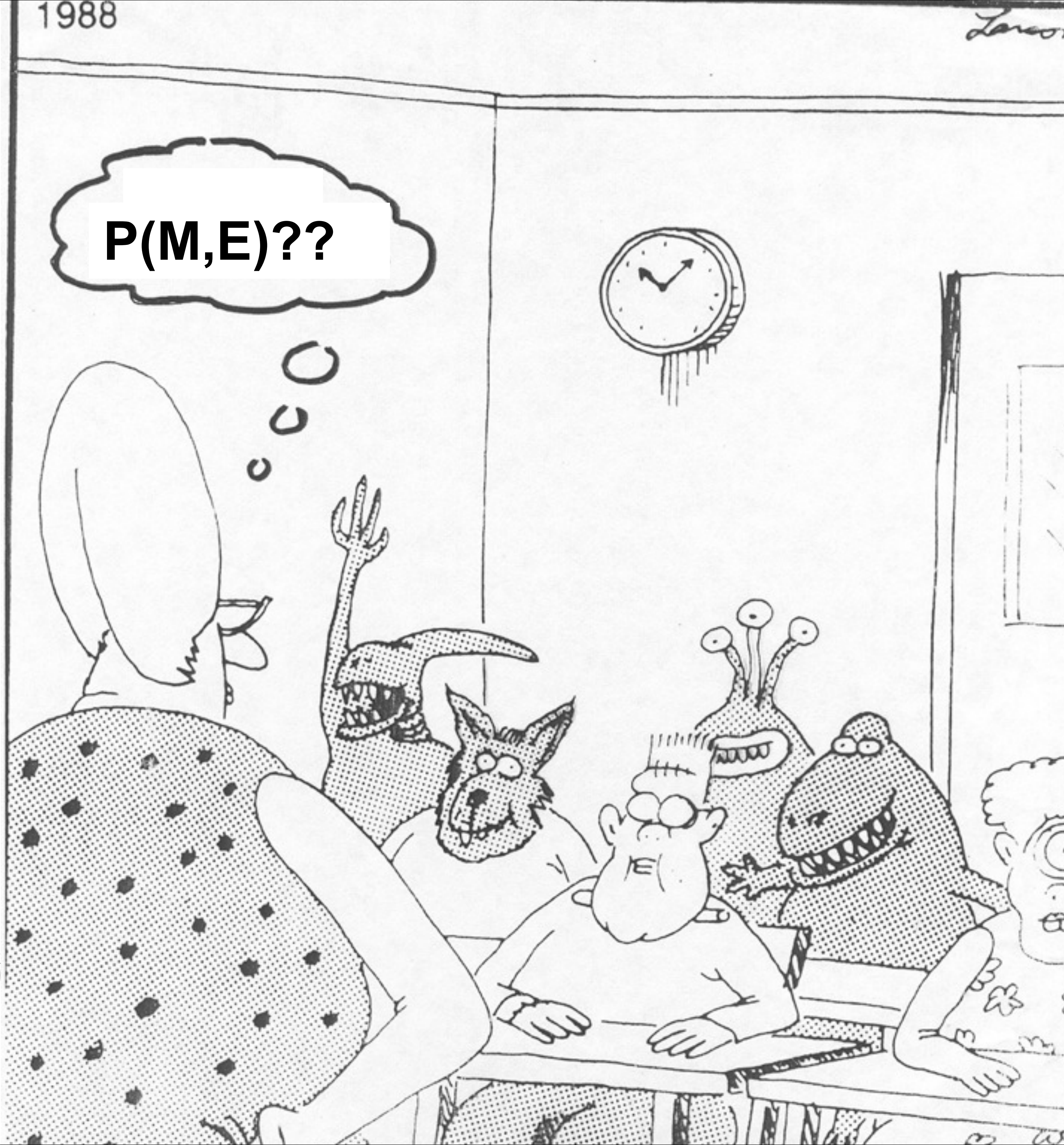
Is this (M+E) possible?

A simple mathematical formulation for Presence follows a Bayesian model of the probability of observing a set of events given a model of reality M:

$$P(E, M) = P(\mathrm{E}|M) \cdot P(M). \tag{11}$$

- *Consistency* of new data with a given model $=$ $P(\text{Events}|M)$ $=$ *Evidence*
- *Possibility*: consistency with already established models $=$ $P(M)$ $=$ *Prior*




1988
P(M,E)??

M-PI and *M*-Psi

- "Does the intended model M_X feel real?"

To be more precise *we can define the M-Place Illusion generated by a VR system with underlying model M_X to be*

$$\mathrm{PI}[M_X] \equiv P(M_X^l, E) = P(E|M_X^l)\, P(M_X^l), \qquad (12)$$

where M_X^l refers to the reality model low level (e.g., perceptual) aspects.

Similarly, the Plausibility Psi will be proportional to PI, to the evidence for higher order models, and to their prior. To be more precise *we define the M-Plausibility of a model M_X given a set of events E to be*

$$\begin{aligned}\mathrm{Psi}[M_X] &\equiv P(M_X, E) = P(M_X^l, E) \cdot P(M_X^h, E) \\ &= \mathrm{PI}[M_X] \cdot P(E|M_X^h) \cdot P(M_X^h|E_{old}) \qquad (13)\end{aligned}$$

Starlab®

Plausibility of Events

- "Is there a model I can construct where this is plausible?"

$$\mathrm{Psi} = \max_M P(M, E)$$

- "Is this really happening??"

$$\mathrm{Psi}_T = \sum_M P(M, E) = P(E)$$

$$\mathrm{Psi}_T \propto \sum_M 2^{-l(M)} \propto 2^{-K_M}$$

- Keep your VR experience "simple"!

Starlab®

RARE

- Let RARE (Real Actions in Real Environments) be an environment with underlying model M_W intended to fool the subject into believing they have travelled in time and space to the Far West.
- Does M_W feel real?

$$PI[M_W] = P(E|M_W^l)P(M_W^l) = 1$$

$$\mathrm{Psi}[M_W] = \mathrm{PI}[M_W] \cdot P(E|M_W^h) \cdot P(M_W^h|E_{old})$$

"You have till sundown to get rid of those awful curtains."

- Is there a model?

$$\mathrm{Psi} = \max_M P(M, E)$$

Is this at all real?

$$\mathrm{Psi}_T = \sum_M P(M, E) = P(E)$$

Other applications

Model building and simplicity

- Robotics
- Education
- Mathematics
- Machine Learning
- Fundamental Physics
- Biology (evolution is computation)

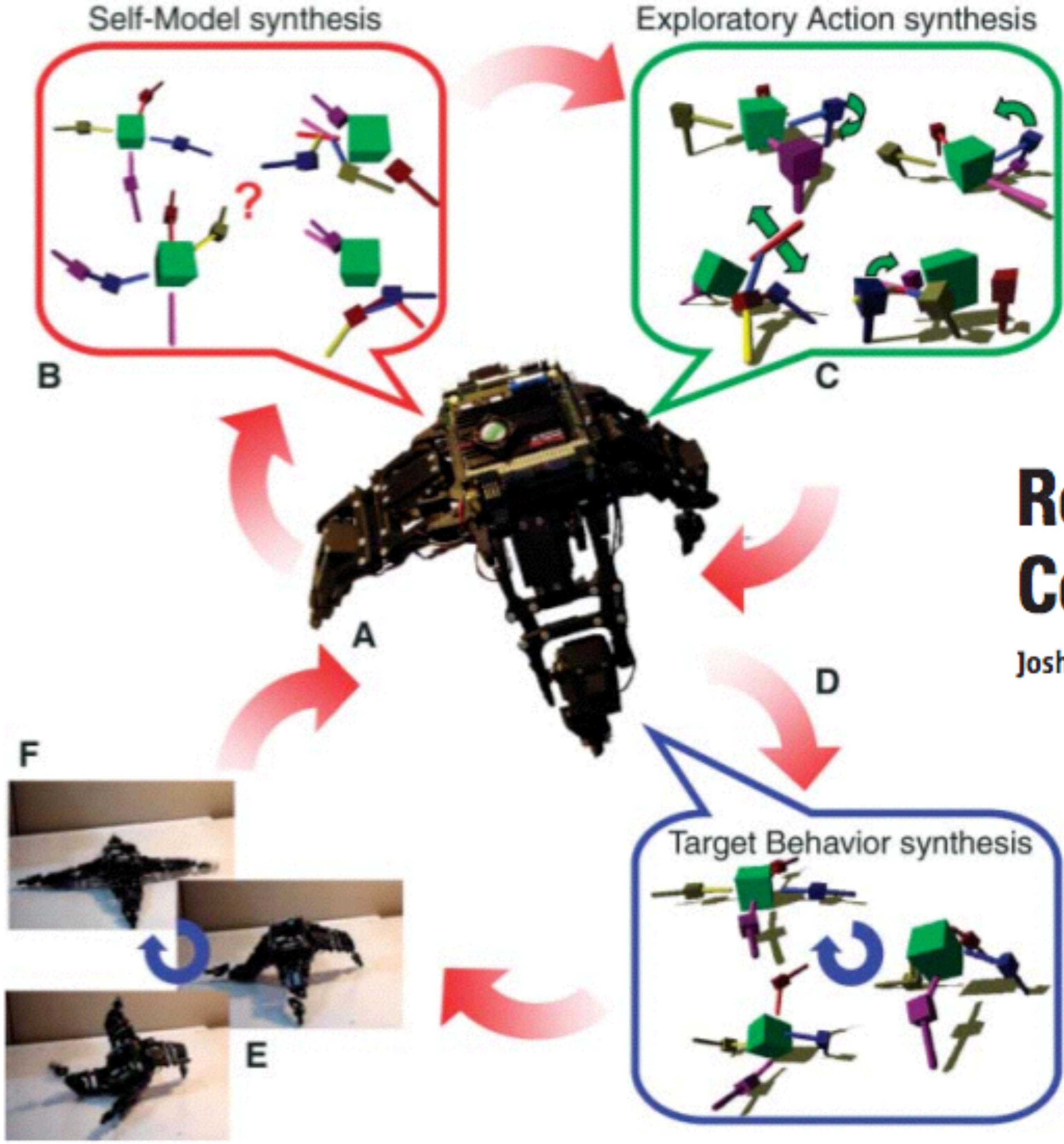


Fig. 1. Outline of the algorithm. The robot continuously cycles through action execution. (**A** and **B**) Self-model synthesis. The robot physically performs an action (A). Initially, this action is random; later, it is the best action found in (C). The robot then generates several self-models to match sensor data collected while performing previous actions (B). It does not know which model is correct. (**C**) Exploratory action synthesis. The robot generates several possible actions that disambiguate competing self-models. (**D**) Target behavior synthesis. After several cycles of (A) to (C), the currently best model is used to generate locomotion sequences through optimization. (**E**) The best locomotion sequence is executed by the physical device. (**F**) The cycle continues at step (B) to further refine models or at step (D) to create new behaviors.

Resilient Machines Through Continuous Self-Modeling

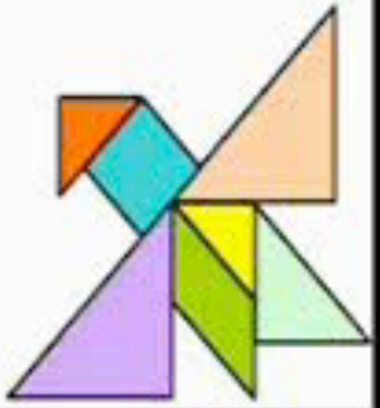
Josh Bongard,[1*†] Victor Zykov,[1] Hod Lipson[1,2]

Cornell University

Robust Machines Through Continuous Self-Modeling

Josh Bongard, Victor Zykov, Hod Lipson

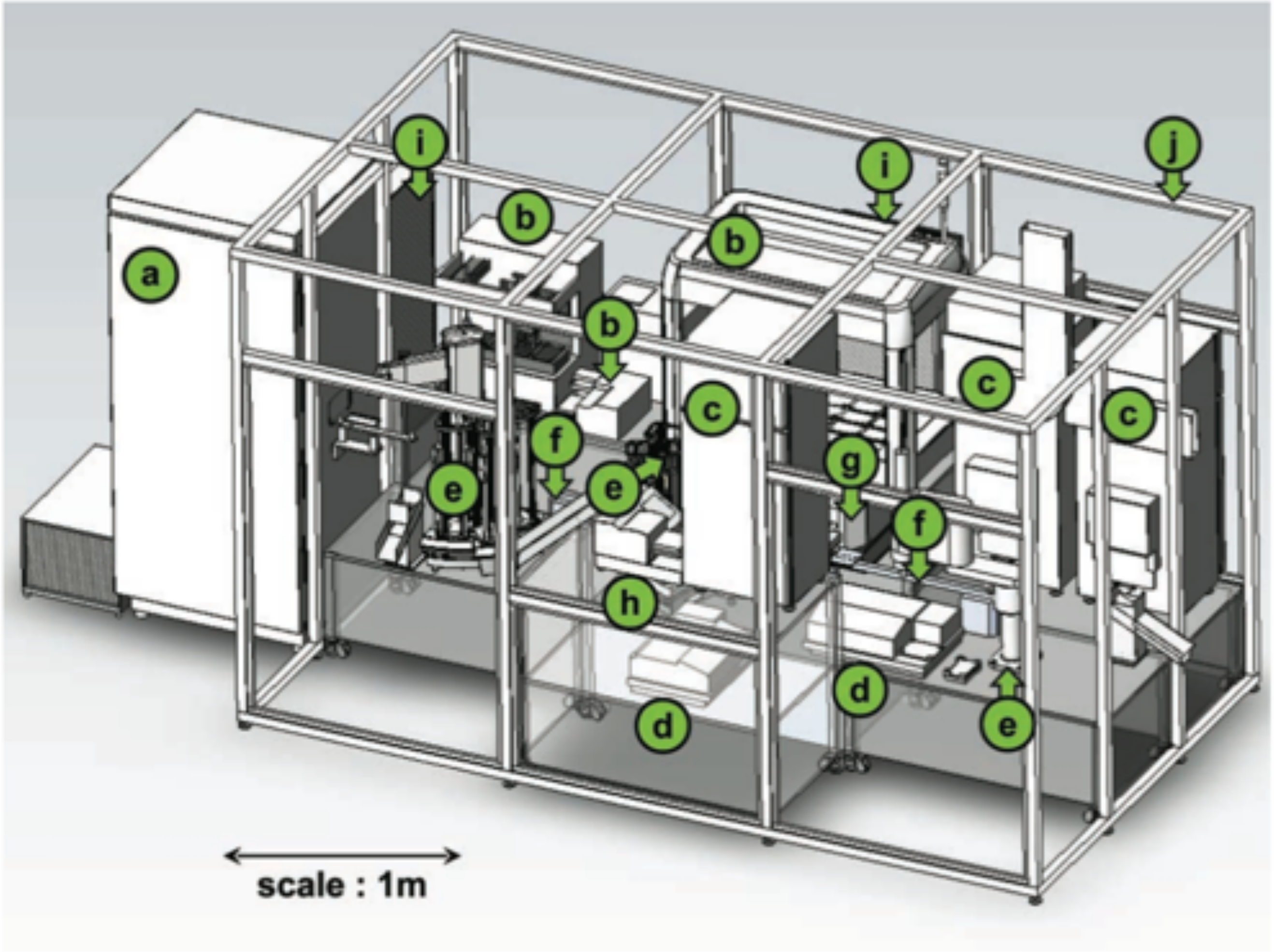
Computational Synthesis Laboratory
Sibley School of Mechanical and Aerospace Engineering
Cornell University

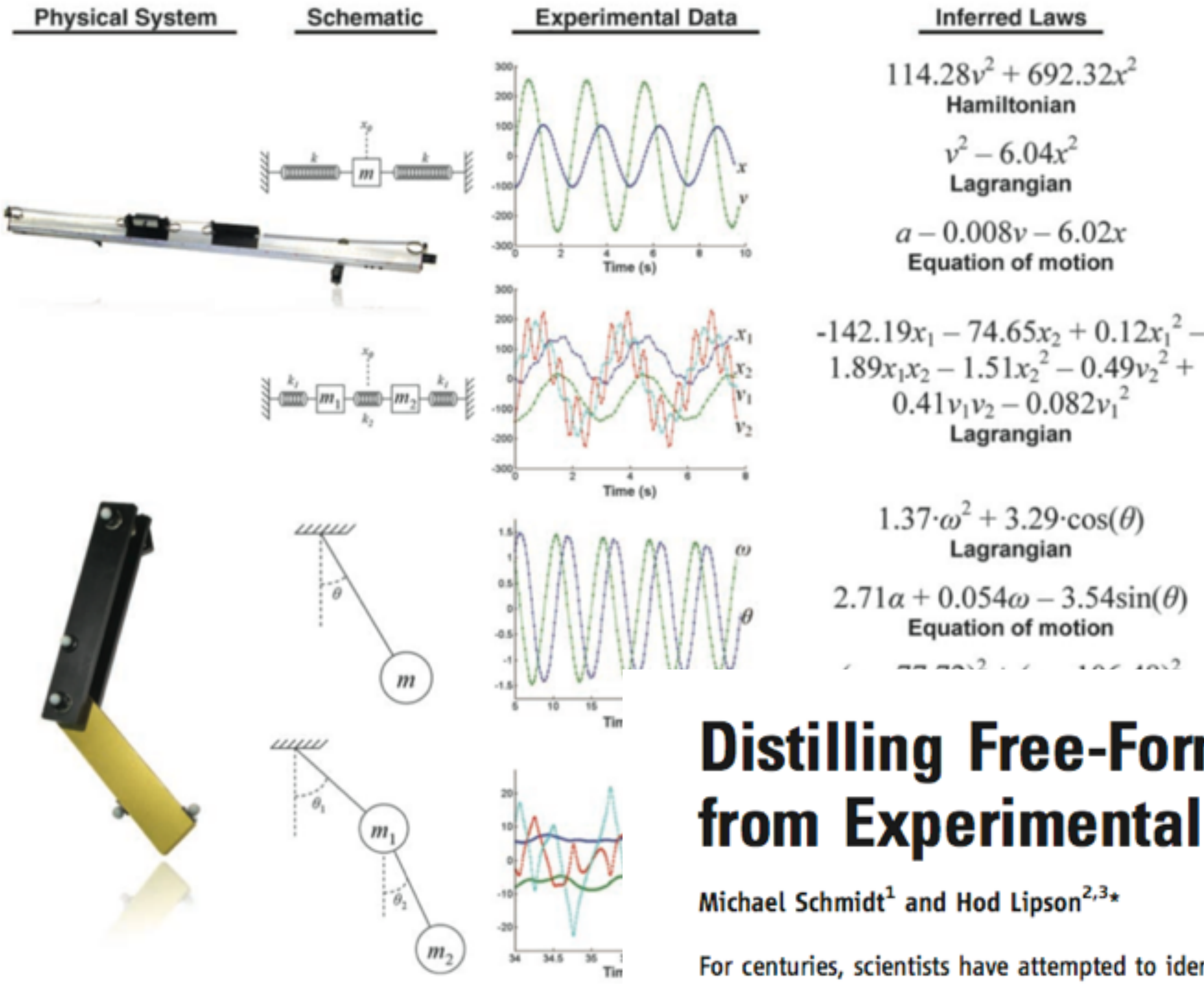


The Automation of Science

Ross D. King,[1]* Jem Rowland,[1] Stephen G. Oliver,[2] Michael Young,[3] Wayne Aubrey,[1]
Emma Byrne,[1] Maria Liakata,[1] Magdalena Markham,[1] Pınar Pir,[2] Larisa N. Soldatova,[1]
Andrew Sparkes,[1] Kenneth E. Whelan,[1] Amanda Clare[1]

The basis of science is the hypothetico-deductive method and the recording of experiments in sufficient detail to enable reproducibility. We report the development of Robot Scientist "Adam," which advances the automation of both. Adam has autonomously generated functional genomics hypotheses about the yeast *Saccharomyces cerevisiae* and experimentally tested these hypotheses by using laboratory automation. We have confirmed Adam's conclusions through manual experiments. To describe Adam's research, we have developed an ontology and logical language. The resulting formalization involves over 10,000 different research units in a nested treelike structure, 10 levels deep, that relates the 6.6 million biomass measurements to their logical description. This formalization describes how a machine contributed to scientific knowledge.





Distilling Free-Form Natural Laws from Experimental Data

Michael Schmidt[1] and Hod Lipson[2,3]*

For centuries, scientists have attempted to identify and document analytical laws that underlie physical phenomena in nature. Despite the prevalence of computing power, the process of finding natural laws and their corresponding equations has resisted automation. A key challenge to finding analytic relations automatically is defining algorithmically what makes a correlation in observed data important and insightful. We propose a principle for the identification of nontriviality. We demonstrated this approach by automatically searching motion-tracking data captured from various physical systems, ranging from simple harmonic oscillators to chaotic double-pendula. Without any prior knowledge about physics, kinematics, or geometry, the algorithm discovered Hamiltonians, Lagrangians, and other laws of geometric and momentum conservation. The discovery rate accelerated as laws found for simpler systems were used to bootstrap explanations for more complex systems, gradually uncovering the "alphabet" used to describe those systems.

The organism as a computer

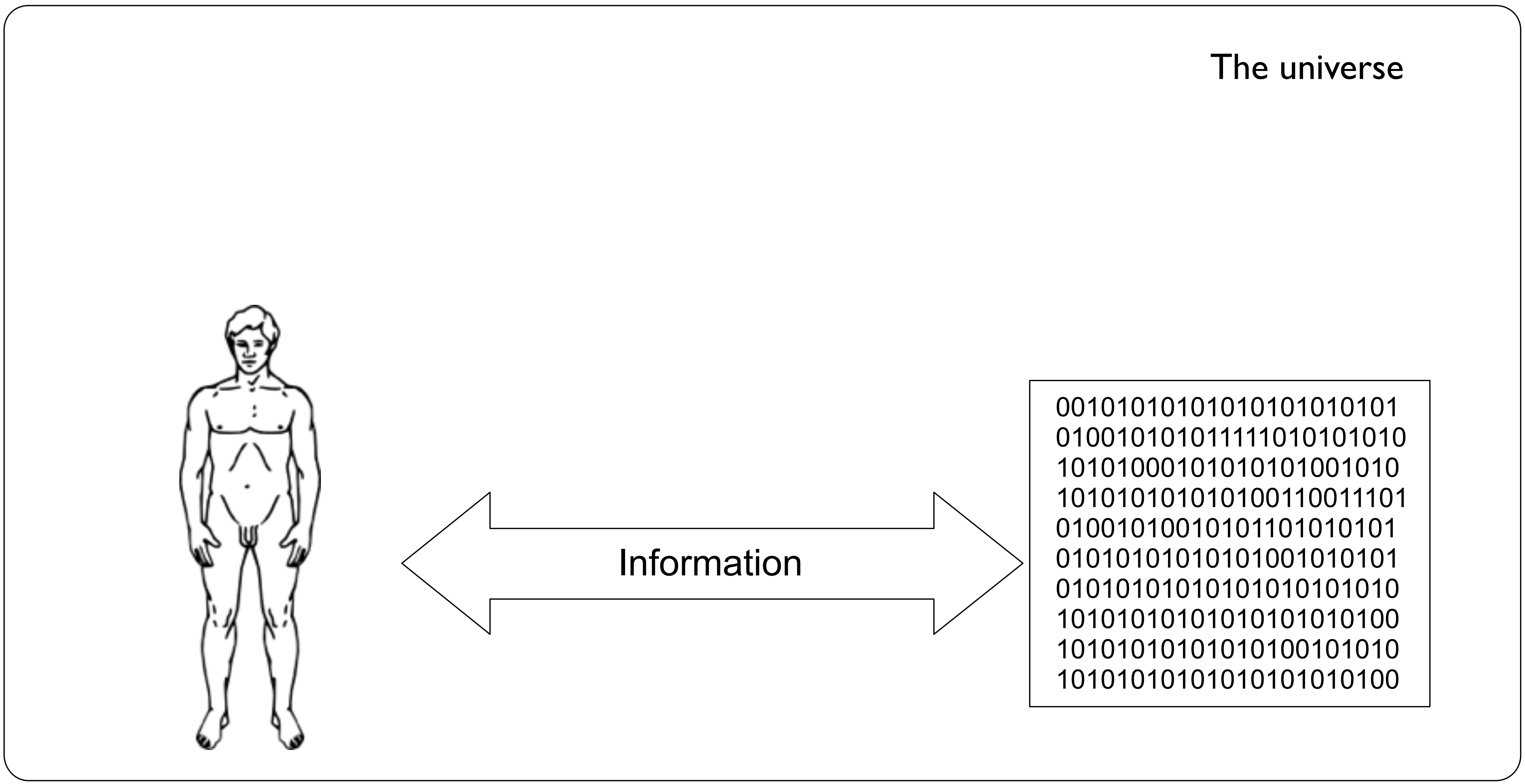
Starlab®

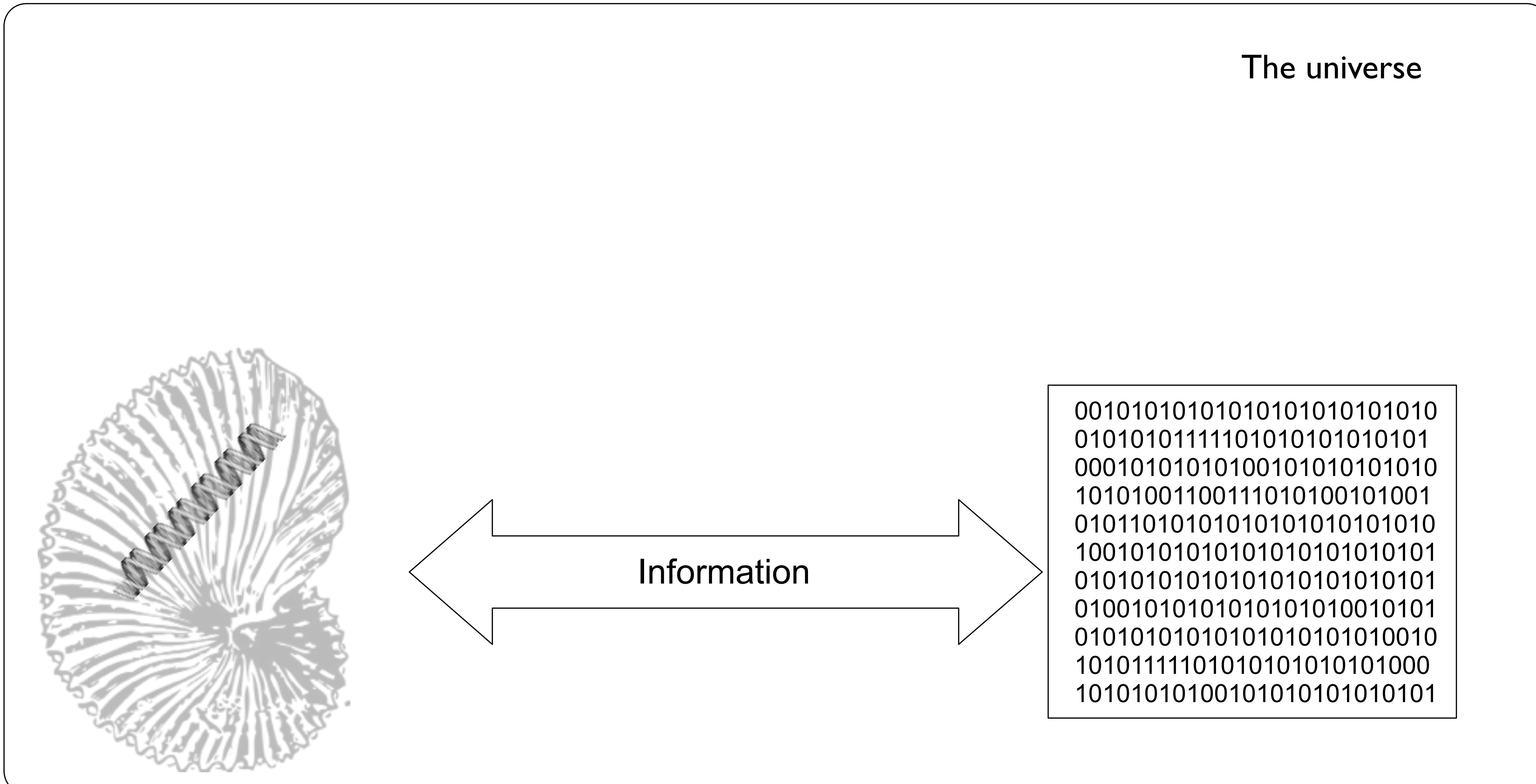
The organic scale

- Organisms can be thought as:
  - Organism genotypes = program or model -- where learning takes place
  - Organism phenotypes = computers running programs
- **Def**: *a living being, or entity or agent, can be defined to be a replicating program that successfully encodes and runs a (partial) model of reality, thus increasing its chances of survival as a replicating program.*
  - Note: some homeostasis is needed for replication
- **Evolution**: the search for better programs. Simplicity may play a role here as well. If so, physiology should find simplicity.

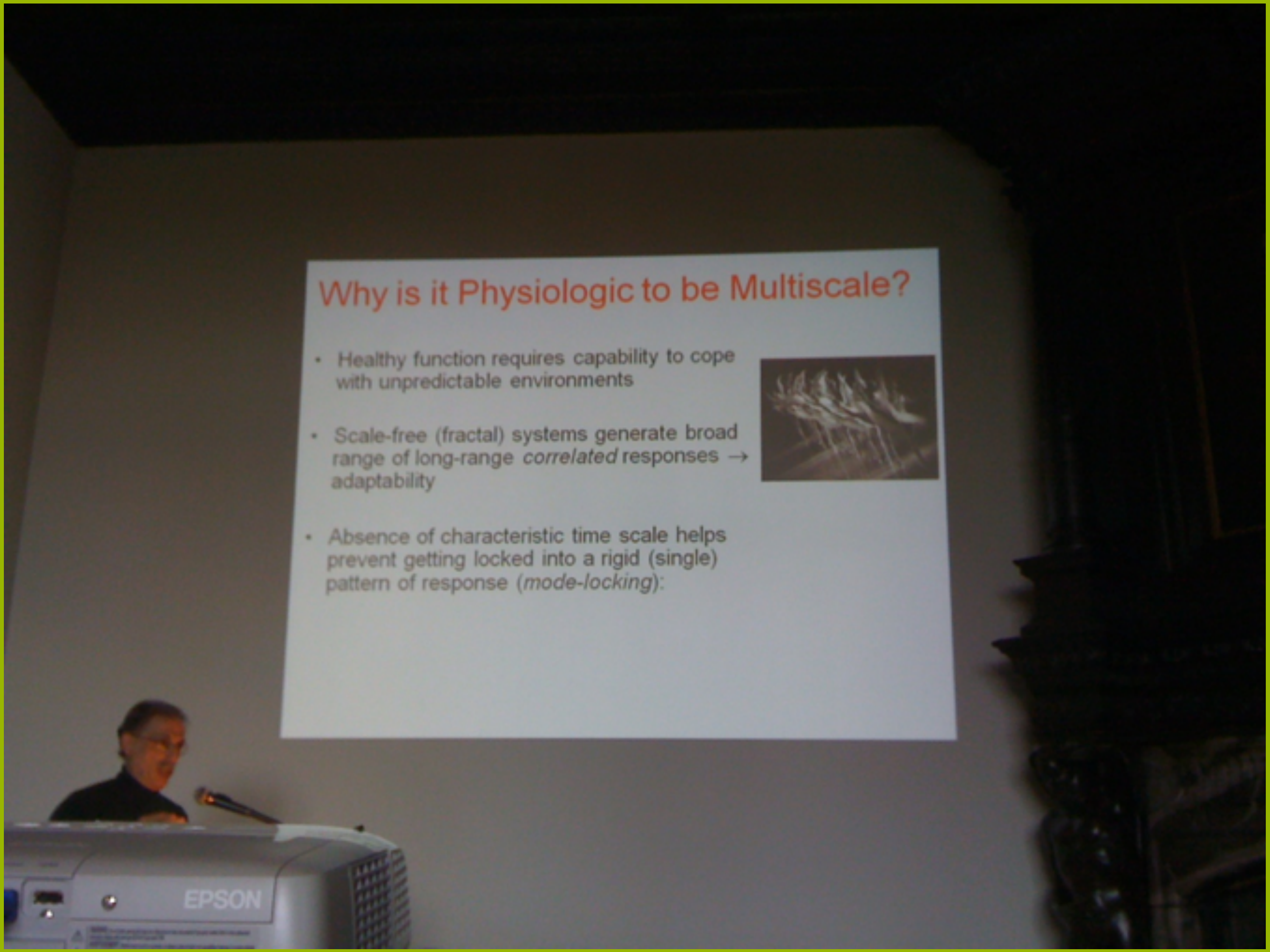
Organisms and information

- Evolution as the search for simple programs (DNA) that replicate well
- Homeostatis is necessary (but not sufficient)



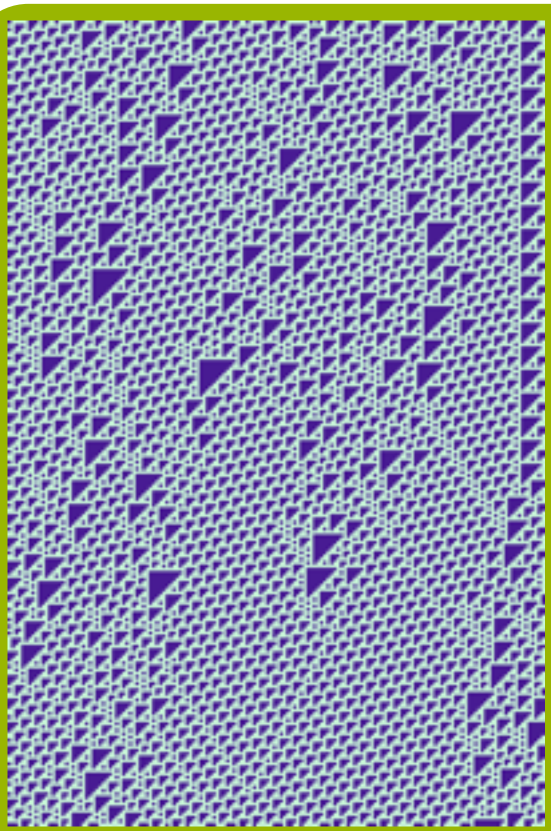

The universe
Information
001010101010101010101010101010
010101011111010101010101010101
000101010101010010101010101010
1010100110011101010100101001
010110101010101010101010101010
100101010101010101010101010101
010101010101010101010101010101
010010101010101010101010010101
010101010101010101010101010010
101011111010101010101010101000
101010101010010101010101010101

Modeling and mutual information



Starlab®

Scale invariance: why?



- Modeling + evolution= hierarchies in organisms
- Scale invariance and observed phenomenon in physiological systems
- Are scale invariances in physiological signals, which relate to control and homeostasis (self-organization), a consequence of the process of evolution building hierarchical modeling (aka control) systems?
- Is complexity the result of ... simplicity?
- How can we test this? An evolutionary approach could be interesting.
- Loss of complexity as organisms age ... modeling breakdown

Starlab®

Even bacteria seem to do it...

Received 25 March; accepted 7 May 2009.
Published online 17 June 2009.

ARTICLES

Adaptive prediction of environmental changes by microorganisms

Amir Mitchell[1], Gal H. Romano[2], Bella Groisman[1], Avihu Yona[1], Erez Dekel[3], Martin Kupiec[2], Orna Dahan[1]* & Yitzhak Pilpel[1,4]*

Natural habitats of some microorganisms may fluctuate erratically, whereas others, which are more predictable, offer the opportunity to prepare in advance for the next environmental change. In analogy to classical Pavlovian conditioning, microorganisms may have evolved to anticipate environmental stimuli by adapting to their temporal order of appearance. Here we present evidence for environmental change anticipation in two model microorganisms, *Escherichia coli* and *Saccharomyces cerevisiae*. We show that anticipation is an adaptive trait, because pre-exposure to the stimulus that typically appears early in the ecology improves the organism's fitness when encountered with a second stimulus. Additionally, we observe loss of the conditioned response in *E. coli* strains that were repeatedly exposed in a laboratory evolution experiment only to the first stimulus. Focusing on the molecular level reveals that the natural temporal order of stimuli is embedded in the wiring of the regulatory network—early stimuli pre-induce genes that would be needed for later ones, yet later stimuli only induce genes needed to cope with them. Our work indicates that environmental anticipation is an adaptive trait that was repeatedly selected for during evolution and thus may be ubiquitous in biology.

Conclusions

Starlab®

Discussion

- A neurocentric, subjective approach to cognition proposed where Information is the most fundamental physical concept.
- Evolution and natural selection lead to compressing or modeling systems, including auto-modeling. Modeling is equivalent to compression or the search for simplicity. In brains, or organisms.
- In this sense, reality, the construction of models from information, is equivalent to simplicity (in brains).
- Simplicity can be naturally described by Kolmogorov or algorithmic complexity.
- The discussion is intended to apply to any cognitive system, simple or complex, natural or artificial (e.g., robots).
- Simplicity may be experimentally explored (e.g., MMN)
- DNA can be thought of as such a cognitive system (with a very long time scales). There may be simplicity lurking in physiology. Is this the origin of scale invariance? Hierarchies --- recursivity. An evolutive approach useful.

For more info Google "Reality as Simplicity" -- http://arxiv.org/abs/0903.1193

Starlab®

Some relevant quotes

- *Reality is merely an illusion, albeit a very persistent one*. [Einstein]
- *Everything should be made as simple as possible, but not simpler*. [Einstein]
- *Pluralitas non est ponenda sine neccesitate* [Ockham, 14th AC]
  models should be no more complex than sufficient to explain the data
- *As for the simplicity of the ways of God, this holds properly with respect to his means, as opposed to the variety, richness, and abundance, which holds with respect to his ends or effects* [Leibniz, 1686]
- *Omnibus ex nihil ducendis sufficit unum* [Leibniz]

Starlab®

Thank you