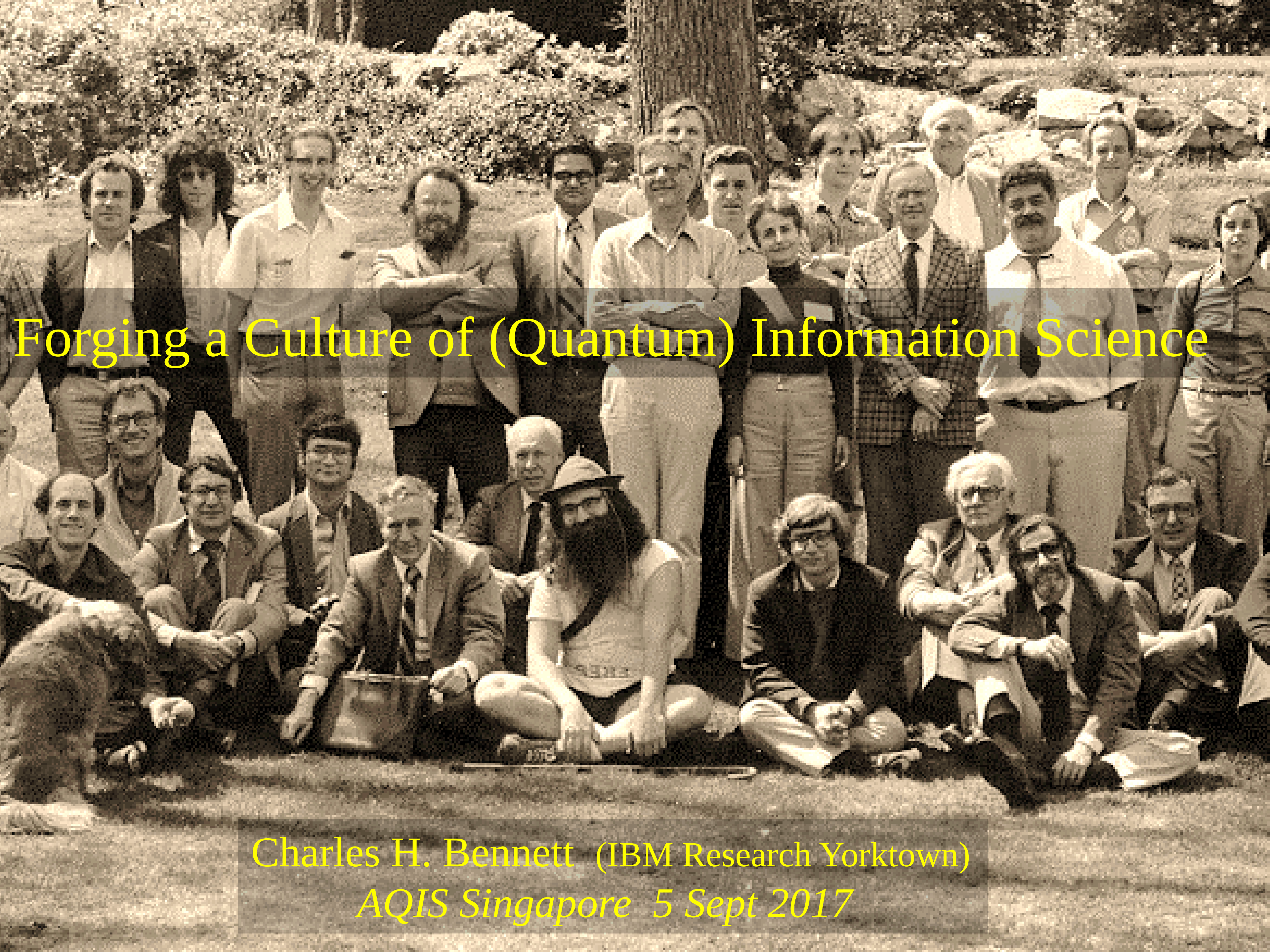


Forging a Culture of (Quantum) Information Science

Charles H. Bennett (IBM Research Yorktown)

AQIS Singapore 5 Sept 2017

Physicists, mathematicians and engineers, guided by what has worked well in their respective disciplines, acquire different scientific tastes, different notions of what constitutes an interesting, well-posed problem or an adequate solution. While this has led to some frustrating misunderstandings, it has invigorated the theory of communication and computation, enabling it to outgrow its brilliant but brash beginnings with Turing, Shannon and von Neumann, and develop its own mature scientific taste, adopting and domesticating ideas from thermodynamics and especially quantum mechanics that physicists had mistakenly thought belonged solely to their field.

Theoretical computer scientists, like their counterparts in physics, suffer and benefit from a high level of intellectual machismo. They believe they have some of the biggest brains around, which they need to tackle some of the hardest problems.

Like mathematicians, they prove theorems and doubt the seriousness of those who don't (e.g. physicists like me).

But beginning in the 1960's a few (e.g.Landauer, Wiesner, Feynman, and Deutsch) tried to bring physical ideas into informatics but were not well understood. Brassard was among the first computer scientists to take these ideas seriously.

Since then the productive friction between the cultures of physics, mathematics and engineering has produced more complete theory of information and communication, extending the old theory as subtly and beautifully as complex numbers extend the reals.

Like other parts of mathematics, information science originated as an abstraction from practical experience. Today's information revolution is based on the brilliant abstractions of Turing and Shannon (among others):

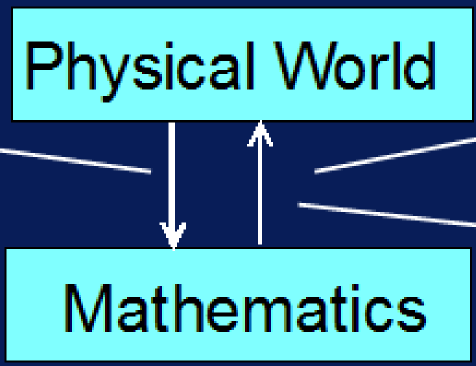
- Turing—a universal, hardware-independent notion of computation
- Shannon—a universal, meaning-independent theory of communication

But now these notions are known to be too narrow.

The subsequent incorporation of two essentially mathematical concepts from physics has led to a more elegant and powerful theory of information and information processing.

Physical (e.g. thermodynamic) resources required for computation

Landauer's slogan "Information is Physical"



Computational resources required to simulate physical states and evolutions

(More mystically, Wheeler's "It from Bit": involvement of information in the creation of physical reality.)

When Turing, Shannon, von Neumann et al formalized the notions of information and computation, they left out a couple of important ideas

Reversibility —— Thermodynamics of Computation

(Superposition —— Quantum Computation)

Conventionally, information carriers have been viewed as what a physicist would call classical systems:

• Their states in principle are reliably distinguishable, and can be observed without disturbing the system.

• To specify the joint state of two non-interacting objects, it suffices to specify the state of each one separately.

But for quantum systems like atoms or photons:

• Attempting to observe a particle's state in general disturbs it, while obtaining only partial information about the state (uncertainty principle).

• Two particles can exist in an *entangled* state, causing them to behave in ways that cannot be explained by supposing that each particle has some state of its own.

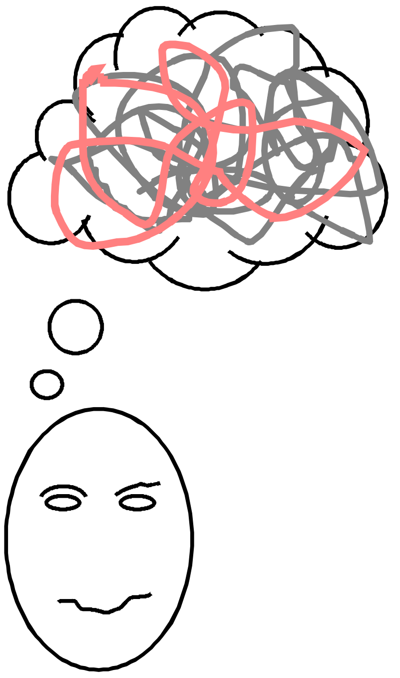
For most of the 20^{th} century, quantum effects in information processing were regarded mainly as a nuisance, because the uncertainty principle makes tiny quantum devices behave less reliably than the classical ideal.

Now it is known that quantum effects also have positive consequences, making possible new kinds of inform-ation processing such as quantum cryptography, and dramatically speeding up some classically hard computations.

These positive consequences are chiefly due to entanglement.

Ordinary classical information, such as one finds in a book, can be copied at will and is not disturbed by reading it.

Quantum information is more like the information in a dream



• Trying to describe your dream changes your memory of it, so eventually you forget the dream and remember only what you've said about it.

• You cannot prove to someone else what you dreamed.

• You can lie about your dream and not get caught.

But unlike dreams, quantum information obeys well-known laws.

Despite the differences there are important similarities between classical and quantum information

All (classical) information is reducible to bits 0 and 1.

All processing of it can be done by simple logic gates (NOT, AND) acting on bits one and two at a time.

Bits and gates are fungible (independent of physical embodiment), making possible Moore's law.

Quantum information is reducible to **qubits** i.e. two-state quantum systems such as a photon's polarization or a spin-1/2 atom.

Quantum information processing is reducible to one- and two-qubit gate operations.

Qubits and quantum gates are fungible among different quantum systems

The central principle of quantum mechanics is

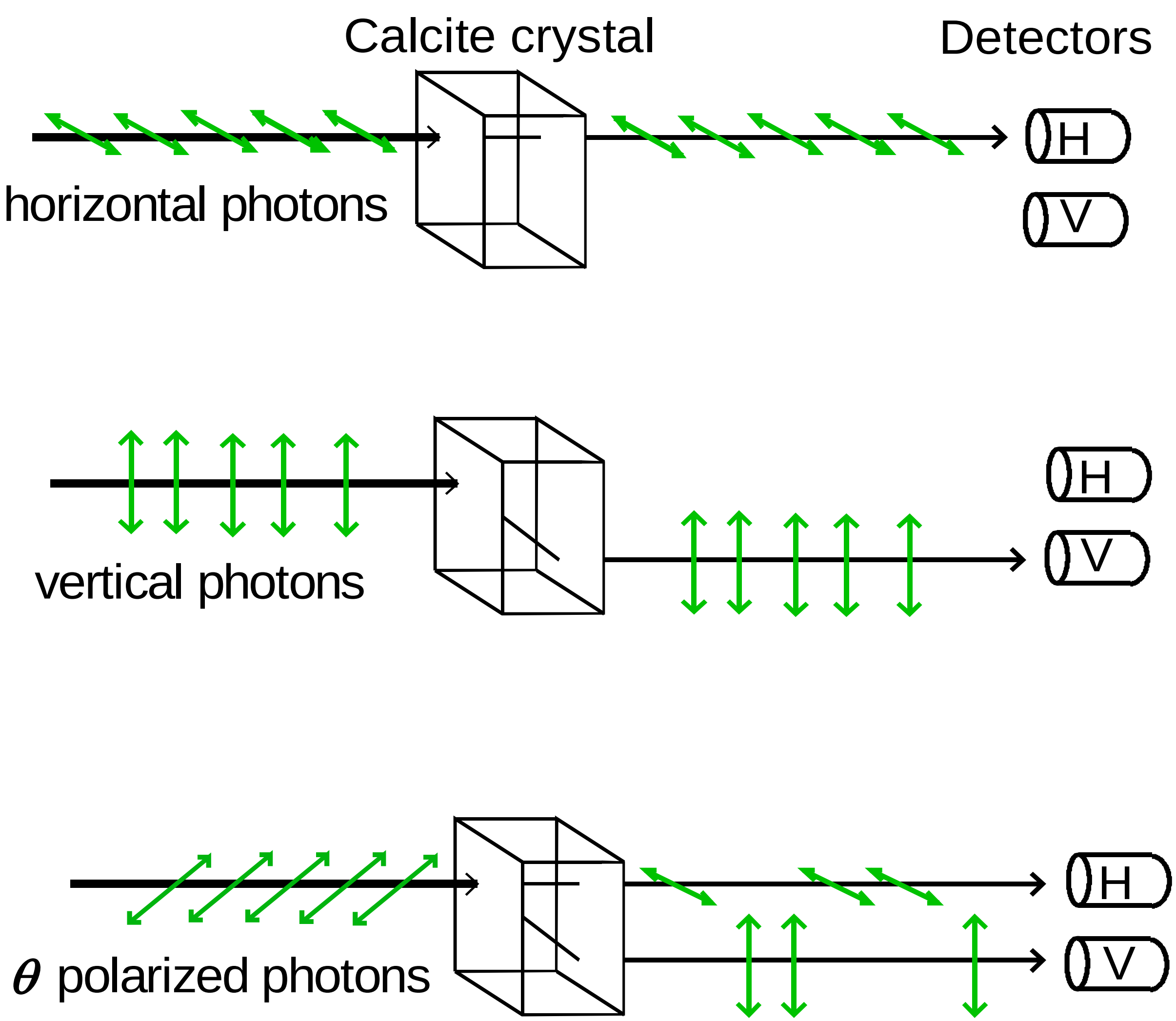
the Superposition Principle:

• Between any two reliably distinguishable states of a physical system (for example vertically and horizontally polarized single photons) there are intermediate states (for example diagonal photons) that are not reliably distinguishable from either original state

• The possible physical states correspond to directions in space—not ordinary 3-dimensional space, but an *n*-dimensional space where *n* is the system's maximum number of reliably distinguishable states.

• Any direction is a possible state, but two states are reliably distinguishable if only if their directions are perpendicular.

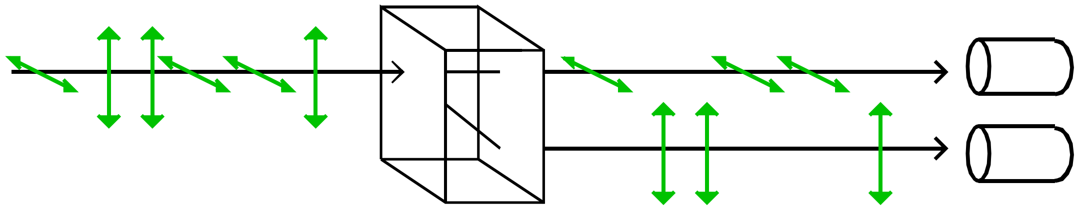
Using Polarized Photons to Carry Information



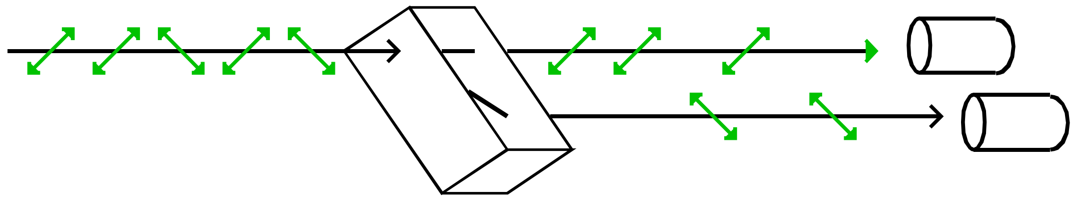
Photons behave reliably if measured along an axis parallel or perpendicular to their original polarization.
Used in this way, each photon can carry one reliable bit of information.

But measuring the photons along any other axis causes them to behave randomly, forgetting their original polarization direction.

A rectilinear (ie vertical vs horizontal) measurement distinguishes vertical and horizontal photons reliably, but randomizes diagonal photons.

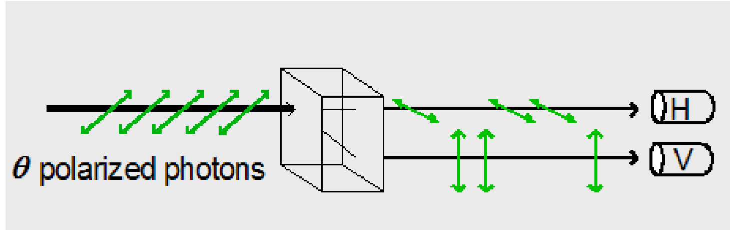


A diagonal measurement distinguishes diagonal photons reliably but randomizes rectilinear photons.



No measurement can distinguish all four kinds. This is not a limitation of particular measuring apparatuses, but a fundamental consequence of the uncertainty principle. This fundamental limitation gives rise to the possibility of quantum money and quantum cryptography.

Bill Wootters' pedagogic analog for quantum measurement



Like a pupil confronting a strict teacher, a quantum system being measured is forced to choose among a set of distinguishable states (here 2) characteristic of the measuring apparatus.

Teacher: Is your polarization vertical or horizontal?

Pupil: Uh, I am polarized at about a 55 degree angle from horizontal.

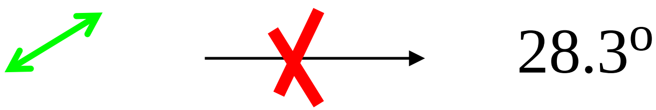
Teacher: **I believe I asked you a question.** Are you vertical or horizontal?

Pupil: Horizontal, sir.

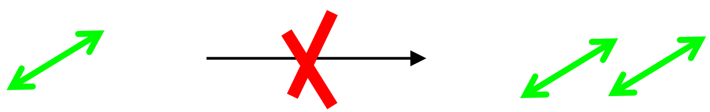
Teacher: Have you ever had any other polarization?

Pupil: No, sir. I was always horizontal.

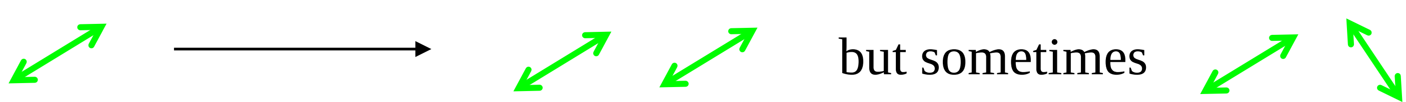
Measuring an unknown photon’s polarization exactly is impossible (no measurement can yield more than 1 bit about it).



Cloning an unknown photon is impossible. (If either cloning or measuring were possible the other would be also).

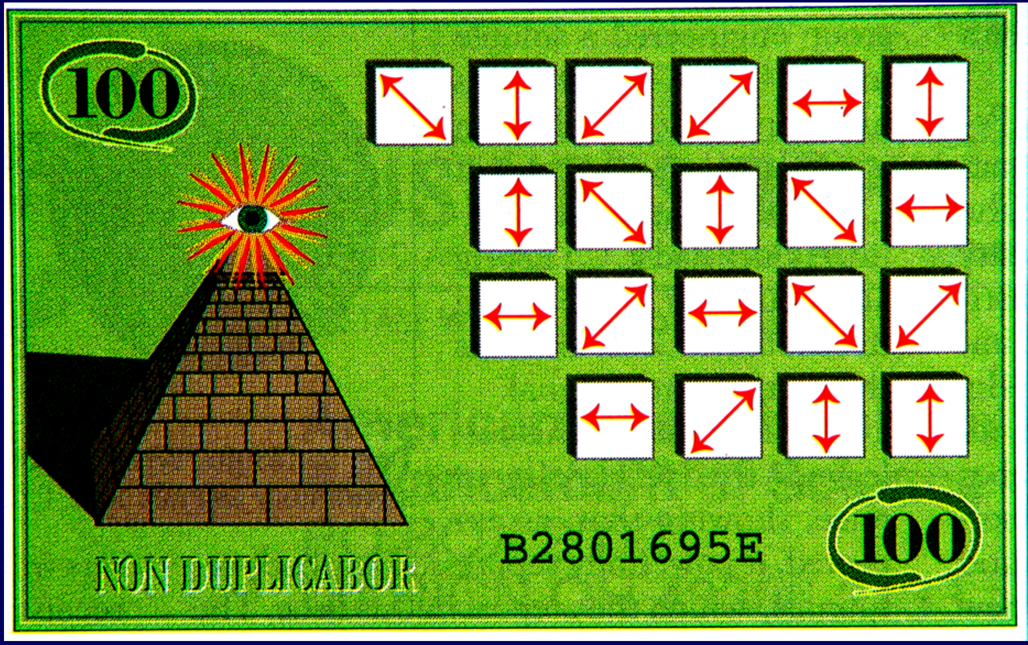


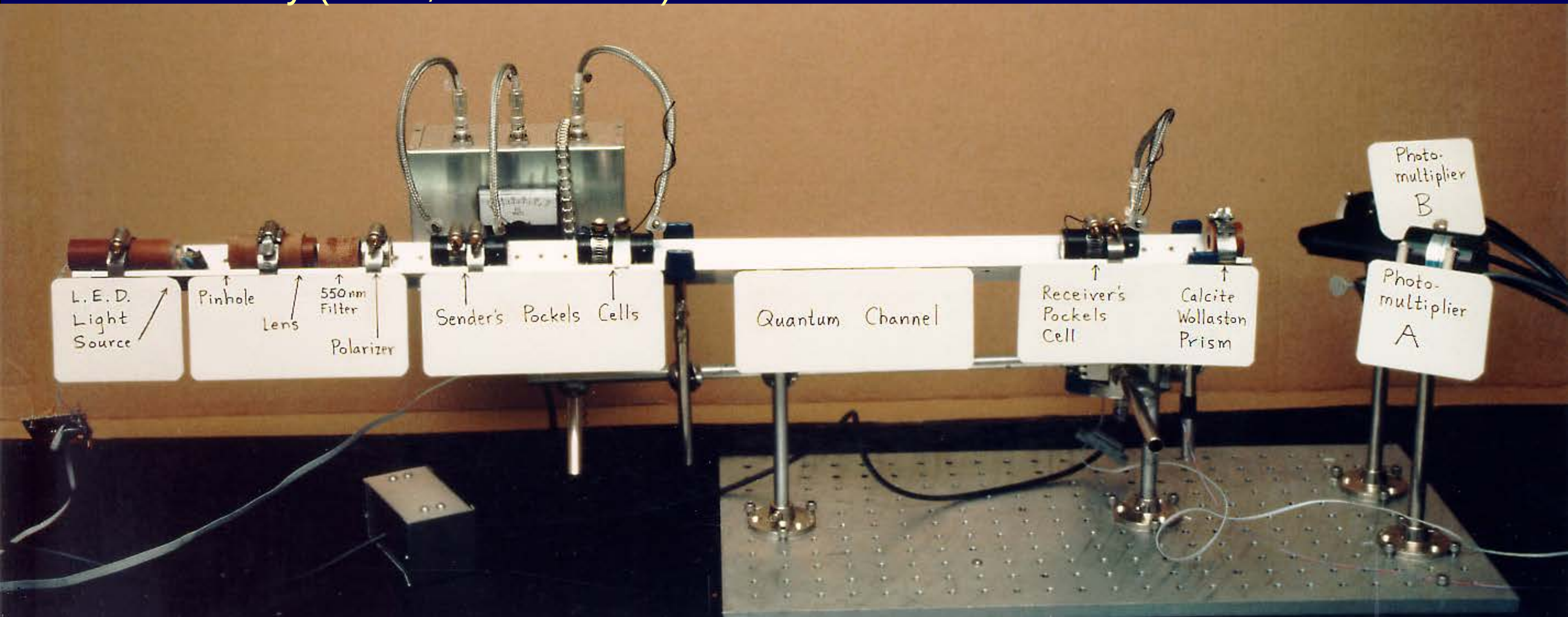
If you try to amplify an unknown photon by sending it into an ideal laser, the output will be polluted by just enough noise (due to spontaneous emission) to be no more useful than the input in figuring out what the original photon’s polarization was.



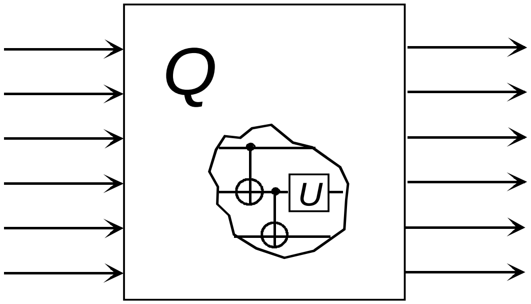
Quantum money (Wiesner '69, '83) cannot be copied by a counterfeiter, but can be checked by the bank, which knows the secret sequence of polarized photons it should contain.

Quantum cryptography uses polarized photons to generate shared secret information between parties who share no secret initially (BB84, BBBSS92…)

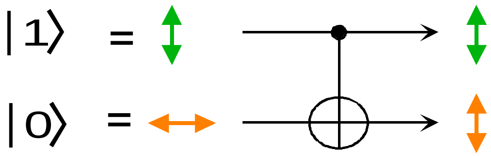


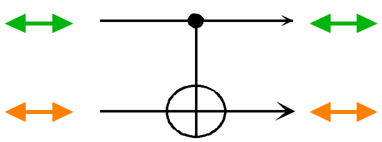


Any quantum data processing can be done by 1- and 2-qubit gates acting on qubits.

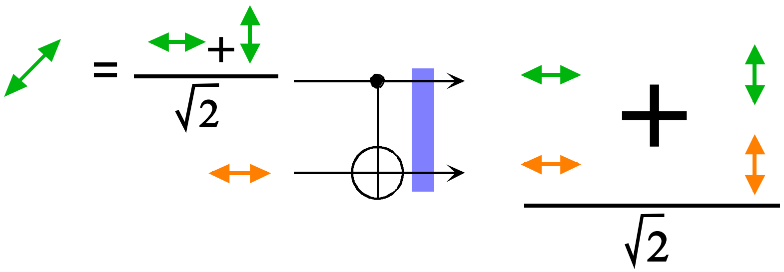


The 2-qubit XOR or "controlled-NOT" gate flips its 2nd input if its first input is 1, otherwise does nothing.



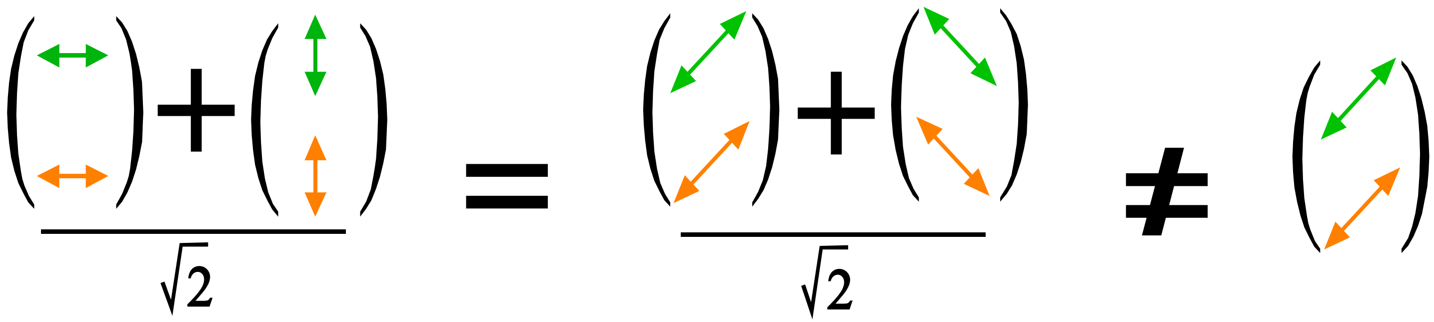


A superposition of inputs gives a superposition of outputs.



An entangled state

This entangled state of two photons behaves in ways that cannot be explained by supposing that each photon has a state of its own.



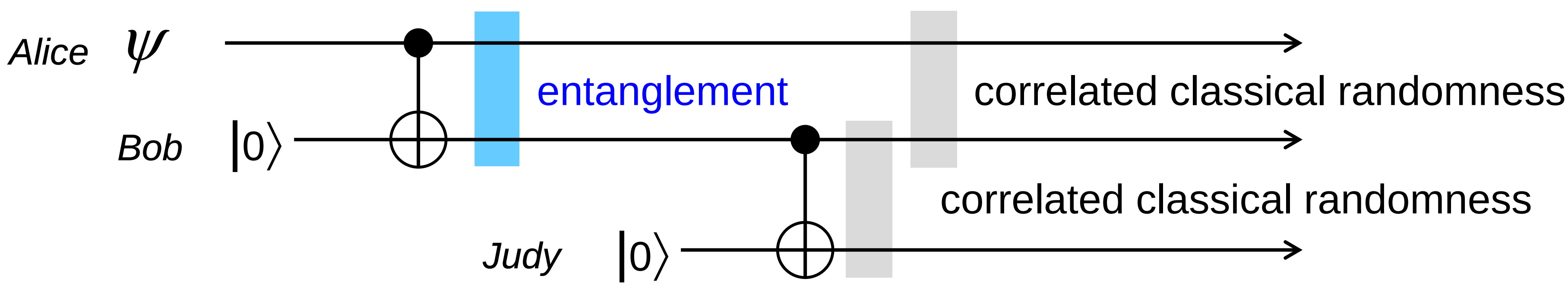
The two photons may be said to be in a definite state of *sameness* of polarization even though neither photon has a polarization of its own.

The Monogamy of Entanglement

• If A and B are maximally entangled with each other, they can't they be entangled with anyone else.

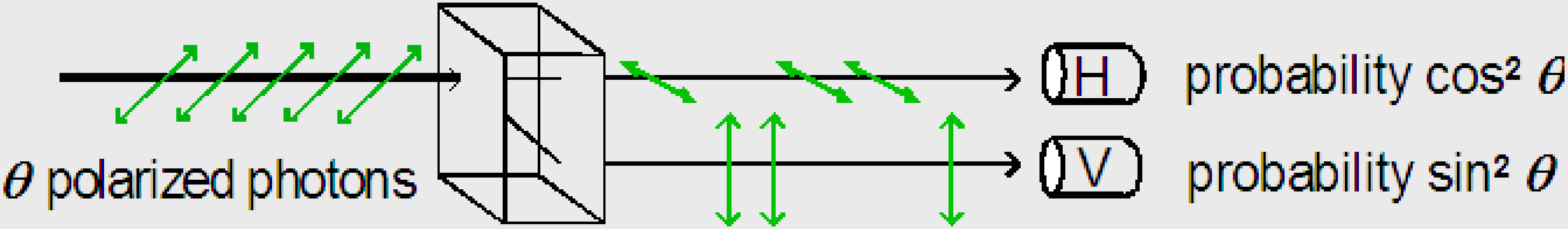
• Indeed classical correlation typically arises from vain attempts to clone entanglement. If one member of an entangled pair tries to share the entanglement with a third party, each pairwise relation is reduced to mere correlated randomness.

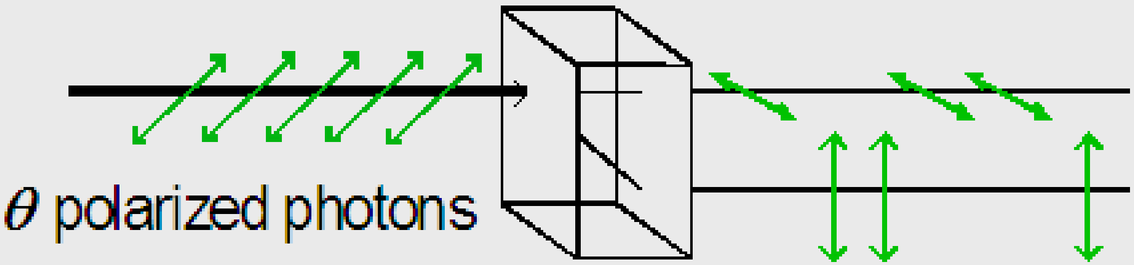
"Two is a couple, three is a crowd."



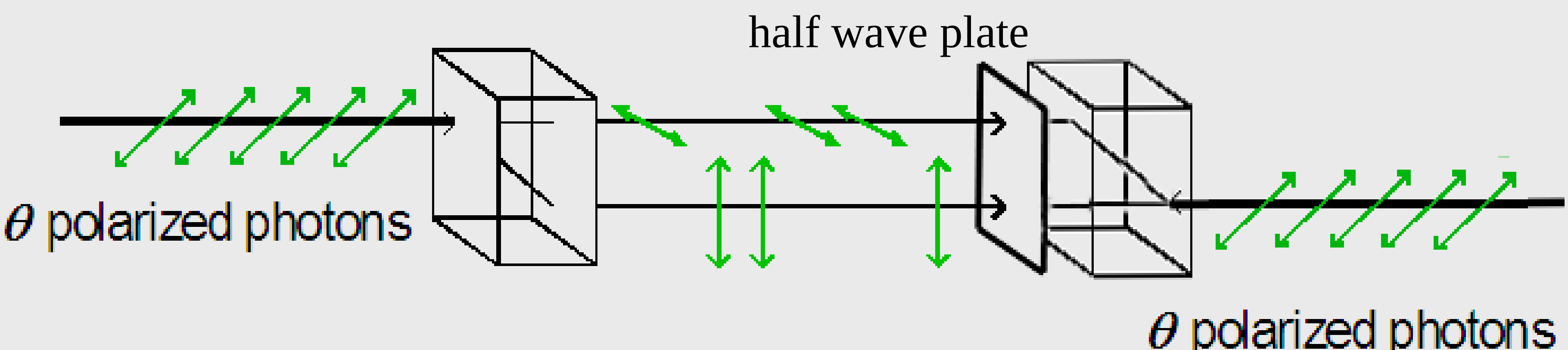
Bob ends up perfectly entangled, not with Alice or Judy, but with the now nontrivial *relationship* between them, an appropriate punishment.

Entanglement and the origin of Quantum Randomness





If no one observes the photons, their random "behavior" can be undone.



Metaphorically speaking, it is the **public embarrassment** of the pupil, in front of the whole class, that makes him forget his original polarization.

Entanglement is ubiquitous: almost every interaction between two systems creates entanglement between them.

Then why wasn't it discovered before the 20th century?

Because of its monogamy.

Most systems in nature, other than tiny ones like photons, interact so strongly with their environment as to become entangled with it almost immediately .

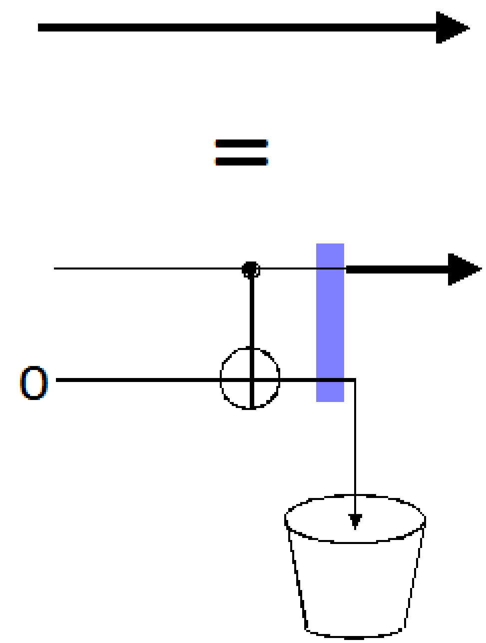
This destroys any previous entanglement that may have existed between internal parts of the system, changing it into mere correlated randomness.

Expressing Classical Data Processing in Quantum Terms

A Classical Bit is a qubit with one of the Boolean values 0 or 1

A classical wire is a quantum channel that conducts 0 and 1 faithfully but randomizes superpositions of 0 and 1.

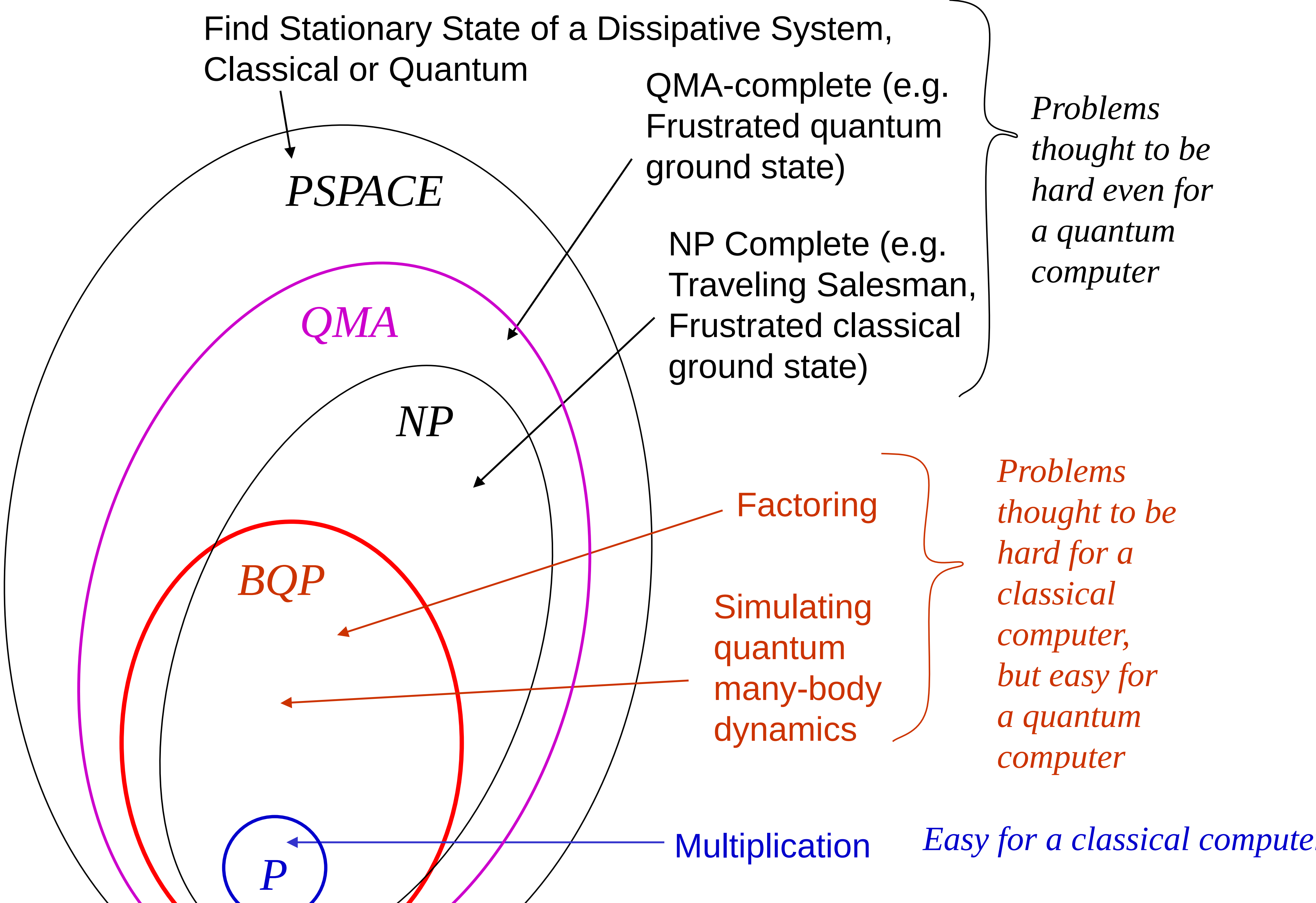
This happens because the data passing through the wire interacts with its environment, causing the environment to acquire a copy of it, if it was 0 or 1, and otherwise become entangled with it.



wastebasket symbolizes loss of
Information into the environment

A classical channel is a quantum channel with an eavesdropper.

A classical computer is a quantum computer handicapped by having eavesdroppers on all its wires.



Result: a thriving quantum theory of info processing

The Einstein -Bohr debate:

When the weird behavior of subatomic particles became evident in the early 20th century, Niels Bohr argued that physicists must learn to accept it. There were two kinds of weird behavior: indeterminacy---the random behavior of individual particles even under completely controlled conditions and entanglement, in which two particles, no matter how far apart, can behave in ways that are individually random, but too strongly correlated for the particles to have been acting independently. Einstein was deeply troubled by these phenomena, disparaging indeterminacy as “God playing dice,” and the entanglement as “spooky action at a distance.” He spent his remaining years searching unsuccessfully for a more naturalistic theory, where every effect would have a nearby cause. Newton’s mechanics, Maxwell’s electromagnetism, and his own relativity share this common-sense property, without which, Einstein thought, science could no longer aspire to be an orderly explanation of nature.

Meanwhile the rest of the physics community, including greats like Schrödinger, Heisenberg, and Dirac, followed Bohr’s advice and accepted these disturbing phenomena, and the mathematics that explained them, as the new normal.

Now, 90 years later, it's pretty clear that the most celebrated scientific mind of the 20th century, flexible enough to bend space and time, still wasn't flexible enough. Quantum randomness and entanglement are real, confirmed by innumerable experiments, and explained in meticulous detail by the theory Einstein disliked. Moreover, quantum theory has played an essential role in technologies such as the laser and the transistor, which could not have been developed on the pre-quantum physics of Newton, Maxwell, and Einstein.

Einstein's mistake was in viewing entanglement as some kind of influence of one particle on the other. The right way to think of it is by giving up basic common sense idea that ~~if the whole is in a perfectly definite state, each part must be in a perfectly definite state.~~ An entangled state is a different kind of state of the whole, which is perfectly definite but requires the parts each to behave randomly. Making any measurement on one of two entangled particles yields a random result, but from that random result, it is possible to perfectly predict what the other particle would do if subjected to the same measurement.

Schrödinger, who understood entanglement better than Einstein, called this effect "steering" but that's a bad name for it. No one would want to drive a car with that kind of steering, because it couples two cars in a way that makes neither one controllable. Both drivers would report that their cars had terrible dangerous steering, so that turning the wheel to the right sometimes caused their car to go right but equally likely caused it to go left. Only afterward, when the drivers compared crash reports, would they realize that their cars had behaved in an eerily correlated way.

Mistakenly believing entanglement could be used for long-range communication, Nick Herbert published a paper and Jack Sarfatti tried to patent this imagined application of it. The refutation of these proposals in the early 1980s, by Dieks, Wootters and Zurek, is part of what led to modern quantum information theory. But this wrong idea, like perpetual motion, is so appealing that it is perpetually being "rediscovered".

A proper understanding of entanglement not only explains why it cannot be used to communicate, but how it brings about the other quantum mystery that troubled Einstein, the random behavior of individual particles. Entanglement's intense correlation is mathematically inseparable from its monogamy, and the random behavior of the parts.

My IBM mentor Rolf Landauer is known for discovering the thermodynamic cost of information erasure, thereby helping launch the theory of reversible computation, many of whose methods proved useful in quantum computation.

With an engineering and physics background, he became concerned with the problem of energy consumption and waste heat removal from computers. The 1981 Endicott conference, which he co-organized with Ed Fredkin and Tom Toffoli of MIT, got the Physics of Computation started as respectable discipline.

Physics of Computation Conference Endicott House MIT May 6-8, 1981

1 Freeman Dyson
2 Gregory Chaitin
3 James Crutchfield
4 Norman Packard
5 Panos Ligomenides
6 Jerome Rothstein
7 Carl Hewitt
8 Norman Hardy
9 Edward Fredkin
10 Tom Toffoli
11 Rolf Landauer
12 John Wheeler
13 Frederick Kantor
14 David Leinweber
15 Konrad Zuse
16 Bernard Zeigler
17 Carl Adam Petri
18 Anatol Holt
19 Roland Vollmar
20 Hans Bremerman
21 Donald Greenspan
22 Markus Buettiker
23 Otto Floberth
24 Robert Lewis
25 Robert Suaya
26 Stan Kugell
27 Bill Gosper
28 Lutz Priese
39 Madhu Gupta
30 Paul Benioff
31 Hans Moravec
32 Ian Richards
33 Marian Pour-El
34 Danny Hillis
35 Arthur Burks
36 John Cocke
37 George Michaels
38 Richard Feynman
39 Laurie Lingham
40 Thiagarajan
41 ?
42 Gerard Vichniac
43 Leonid Levin
44 Lev Levitin
45 Peter Gacs
46 Dan Greenberger

But Landauer had some ideas about mathematics which I think were as unproductive as Einstein's ideas about entanglement.

In blunt opposition to Wheeler's enigmatic and mystical "It from Bit", Landauer's favorite slogan was "Information is Physical." He took this to mean that mathematical concepts incapable of physical embodiment, such as the 2^{1000}th digit of pi, when there are not that many atoms in the universe, were of dubious reality and probably not worth thinking about.

I told him this reminded me of the ancient Greeks' discomfort with infinity and irrational numbers, both concepts that later proved a very fruitful both theoretically and practically.

Sarfatti's and Herbert's ideas about entanglement were so wrong that they facilitated the acceptance of the no-cloning theorem as a central fact about quantum information. The theorem had actually been proved in 1970, by J. L. Park, [Foundations of Physics, 1, 23-33, (1970)], but his paper went unnoticed until the theorem was rediscovered by Dieks and by Wootters and Zurek at a time more ripe for its importance to be appreciated.

Moral: wrong ideas sometimes stimulate scientific progress.

Conversely, as we shall see later, correct ideas—indeed quantum mechanics itself—sometimes retard scientific progress.

Physical (e.g. thermodynamic) resources required for computation

Physical World

Mathematics

Computational resources required to simulate physical states and evolutions

Landauer's slogan "Information is Physical"

(More mystically, Wheeler's "It from Bit": involvement of information in the creation of physical reality.)

When Turing, Shannon, von Neumann et al formalized the notions of information and computation, they left out a couple of important ideas

Reversibility —— Thermodynamics of Computation

(Superposition —— Quantum Computation)

The analogy between computation and physical dynamics is much older. For example Galileo's "The book of nature is written in the language of mathematics" and Laplace's elegant description of a universe governed by Newtonian mechanics,

"We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes."
Pierre Simon Laplace 1814

Note that this computation is deterministic and reversible, a feature seemingly lost with quantum indeterminism, but then recovered in a more inclusive form with unitary quantum evolution.

The Second Law of Thermodynamics has many avatars, manifestations that seem unrelated but in fact are equivalent to one another

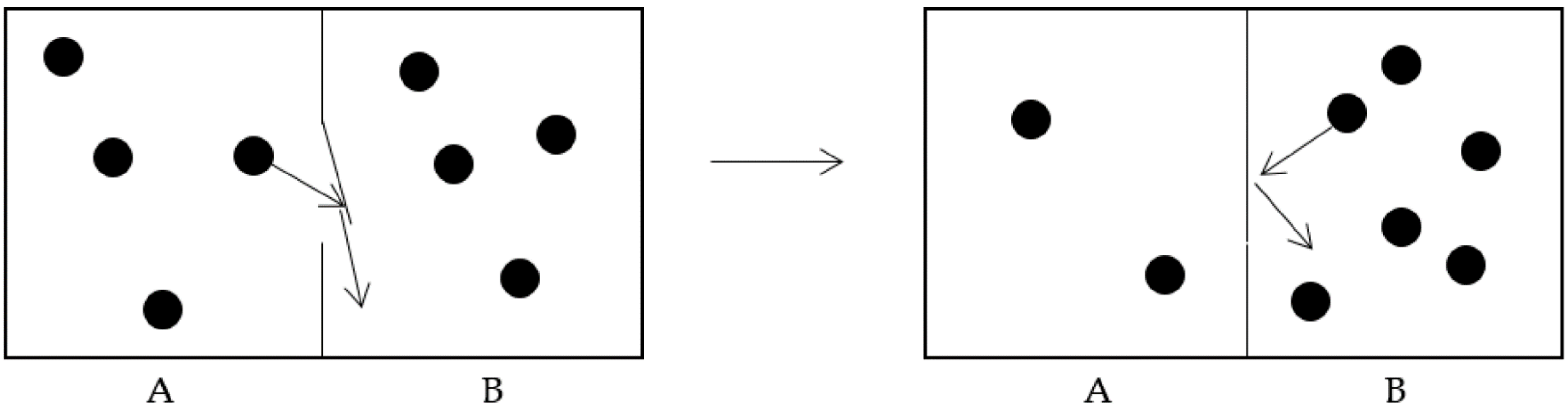
- Heat cannot, of itself, pass from one body to a hotter body. (Kelvin / Flanders & Swann)
- No physical process has as its sole result the conversion of heat into work. (Clausius)
- You can't see anything inside a uniformly hot furnace by the light of its own glow. (Kirchoff)
- Ice Skating (Possible since ice floats, and melts under pressure)
- No physical process has as its sole result the erasure of information. (Landauer / Schumacher)

Maxwell's Demon and its Refutation

If we conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are as essentially finite as our own, would be able to do what is impossible to us. For we have seen that molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower molecules to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics. *James Clerk Maxwell 1867*

Smoluchowski's trap door demon (1912), and his exorcism of it

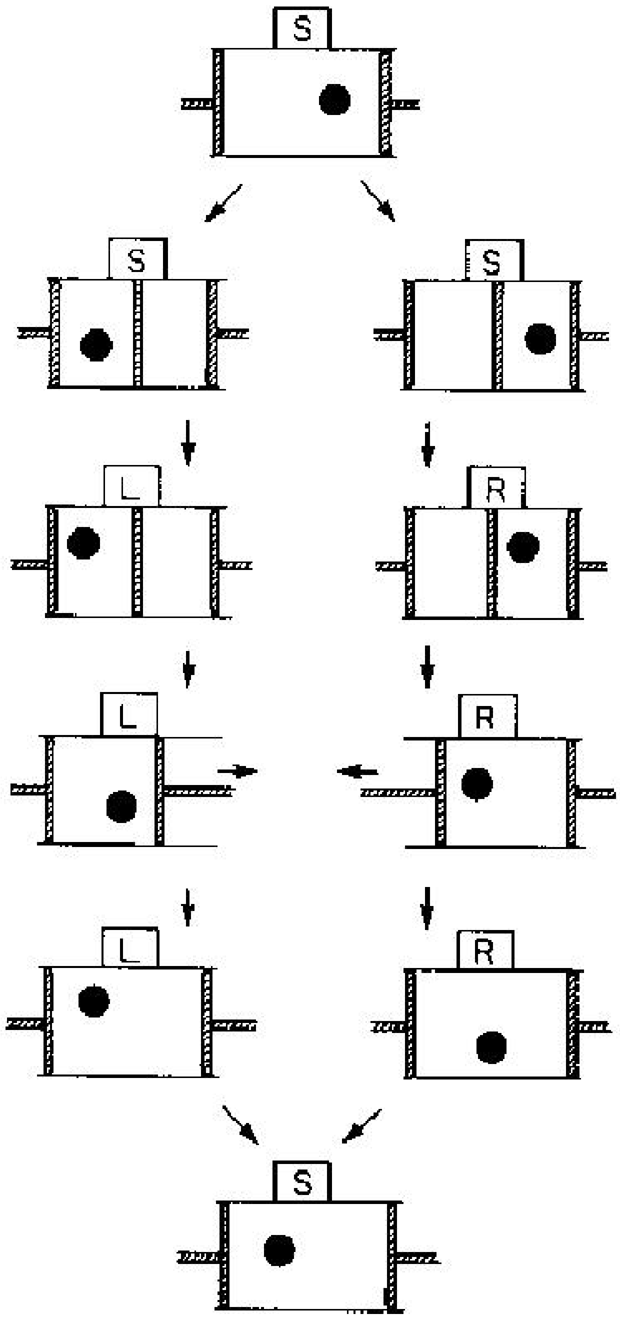
A spring-loaded trap door, light enough to be pushed open by molecular impacts, would seem to violate the Second Law, effortlessly collecting molecules on the right in a pressure version of Maxwell's temperature demon.



But, Smoluchowski argued, if the door were that light and the spring that weak, the door would soon heat up to the same temperature as the gas and undergo random motion of its own, swinging open and shut. It would then swing shut against a molecule that had wandered in front of it, pushing it to the left, just as often as it would be pushed open by a molecule striking it from the left, and there would be no net flow.

Despite Laplace’s deterministic universe, and what Turing and Gödel were about to discover about computation, early 20th century physicists were reluctant to think of ***thought*** itself as a mechanistic process, so Smoluchowski's neat exorcisim of the demon unravelled somewhat in subsequent decades. Leo Szilard’s 1929 paper, in which he introduced his famous engine, was titled *“On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings.”*

The situation was further muddied by the discovery of quantum mechanics, which problematized the previously uncontroversial act of measurement. This tempted physicists to look for an irreducible cost of *information acquisition*, *transmission* or *processing*, when they would have done better to think like Smoluchowski. Even von Neumann incorrectly asserted in a 1949 lecture that each elementary act of information, each decision of a two-way alternative or transmission of a bit of information, must have a thermodynamic cost of kT ln 2 at temperature T. In 1961 Rolf Landauer correctly identified ***information destruction*** as the fundamentally costly act.



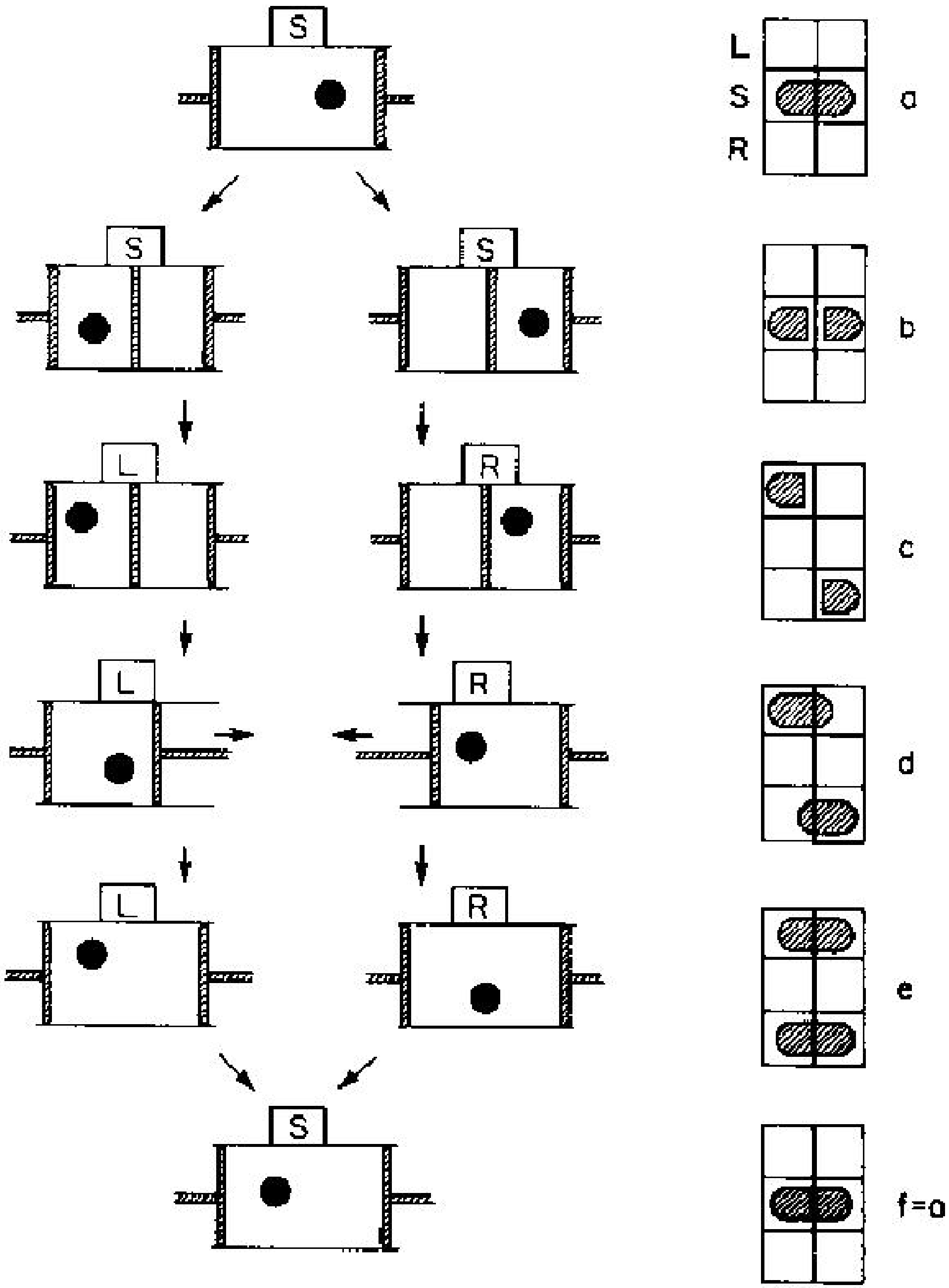
Sziliard's 1929 Engine, attempting to repeatedly extract isothermal work from a molecule.

Demon inserts partition in middle, trapping the molecule on one side or the other.

Measures and remembers which side molecule is on.

Inserts piston on opposite side, removes partition, then lets molecule do kT ln 2 of isothermal work pushing piston back to its original position.

Finally demon resets its memory and repeats the cycle

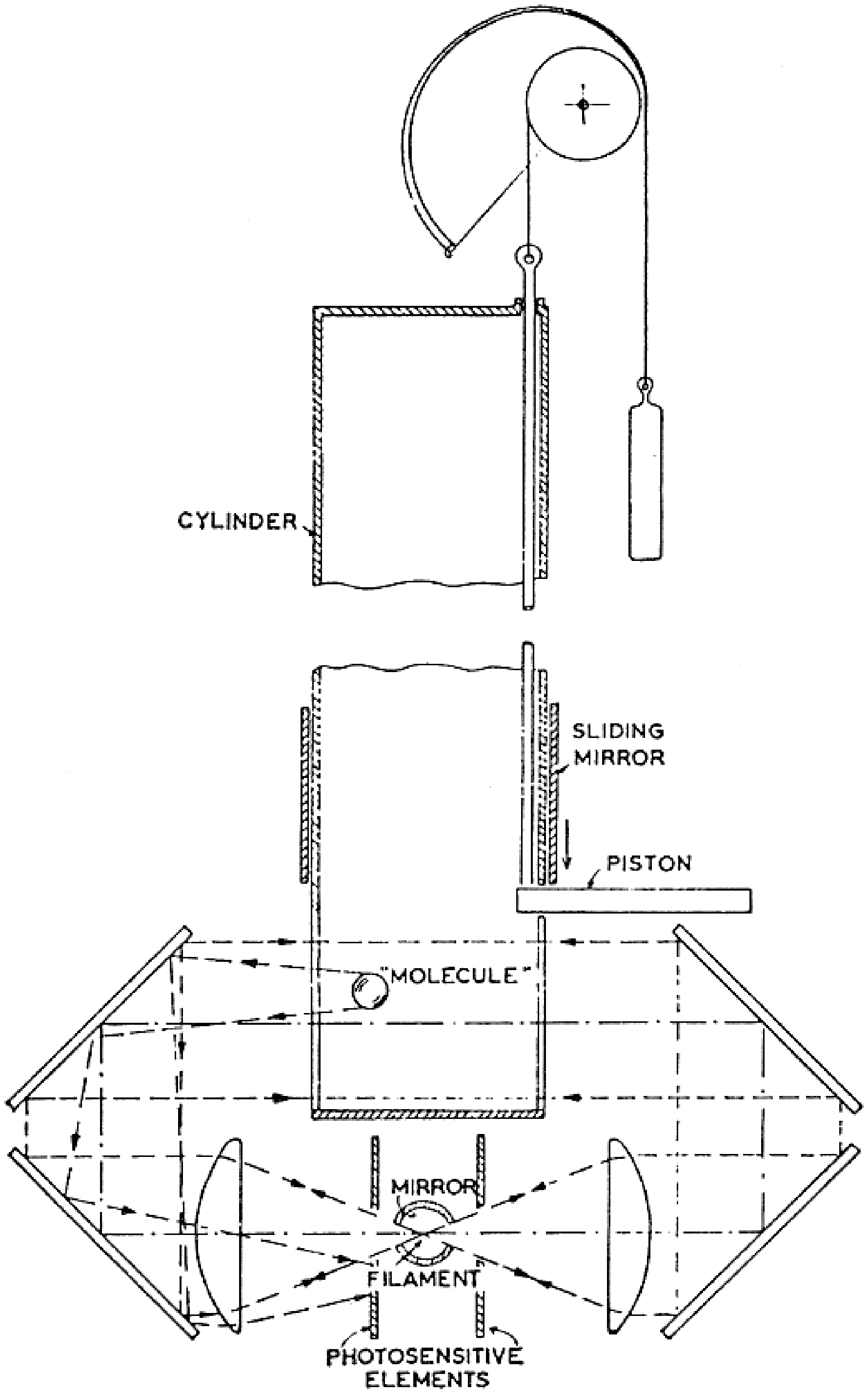


Joint phase space diagram of molecule and memory register shows how, if the register is initially in a standard blank state S, the measurement can be done reversibly, but the final step (f) of resetting the memory entails a compression of phase space that must pay back all the work gained step (d).

Szilard’s 1929 paper made this clear in its equations, but unfortunately not its prose, so the notion that measurement is intrinsically irreversible persisted.

Examples of that sloppy thinking due to misapplication of quantum mechanics to Maxwell’s demon include Leon Brillouin's 1956 argument that to even see a molecule, against the background of quantum black body radiation at temperature T, a demon would need to expend at least one photon more energetic than kT.

Denis Gabor's 1961 refutation of his own high-compression version of Szilard’s engine was the most intricately unnecessary invocation of quantum optics to prove what Smoluchowski had already proved.



Denis Gabor's high-compression Szilard engine (1961).

- Light beam circulates losslessly across one end of a long cylinder
- Photosensors detect when molecule wanders into the beam, and insert a piston to trap it there.
- Piston extracts $kT \ln (V/V_0)$ work by a very long isothermal power stroke.
- Some of the work is used to reset piston & recreate the light beam.
- Since it takes only a fixed amount of work **w** to do that, one can break the Second Law by making V so large that $kT \ln(V/V_0) > \mathbf{w}$.

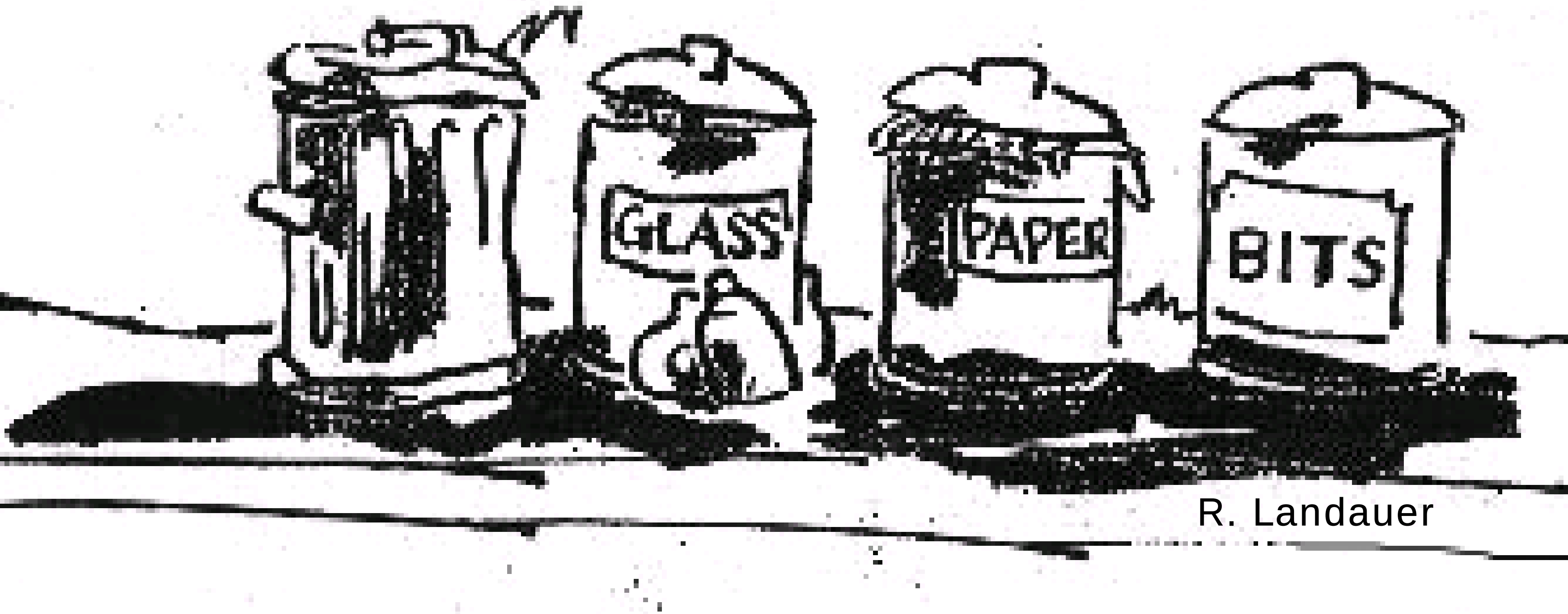
What keeps it from breaking the 2nd Law?

Can you guess Gabor's answer? (hard)

Can you guess the correct answer? (easy)

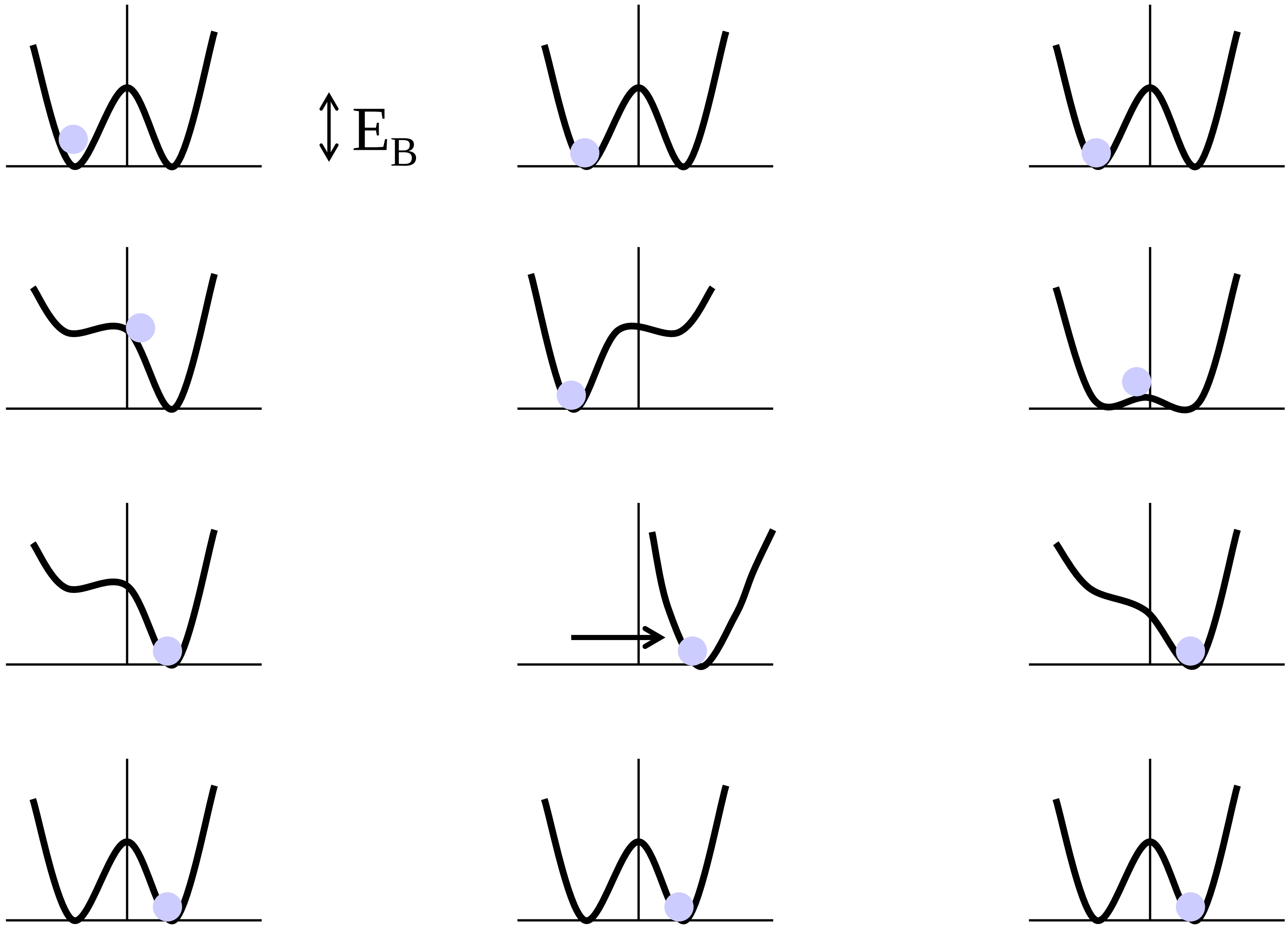
*See answers at end of talk

Another retarding idea was that information was a valuable resource. Landauer showed that unwanted information is actually a waste product, requiring work to get rid of.



R. Landauer

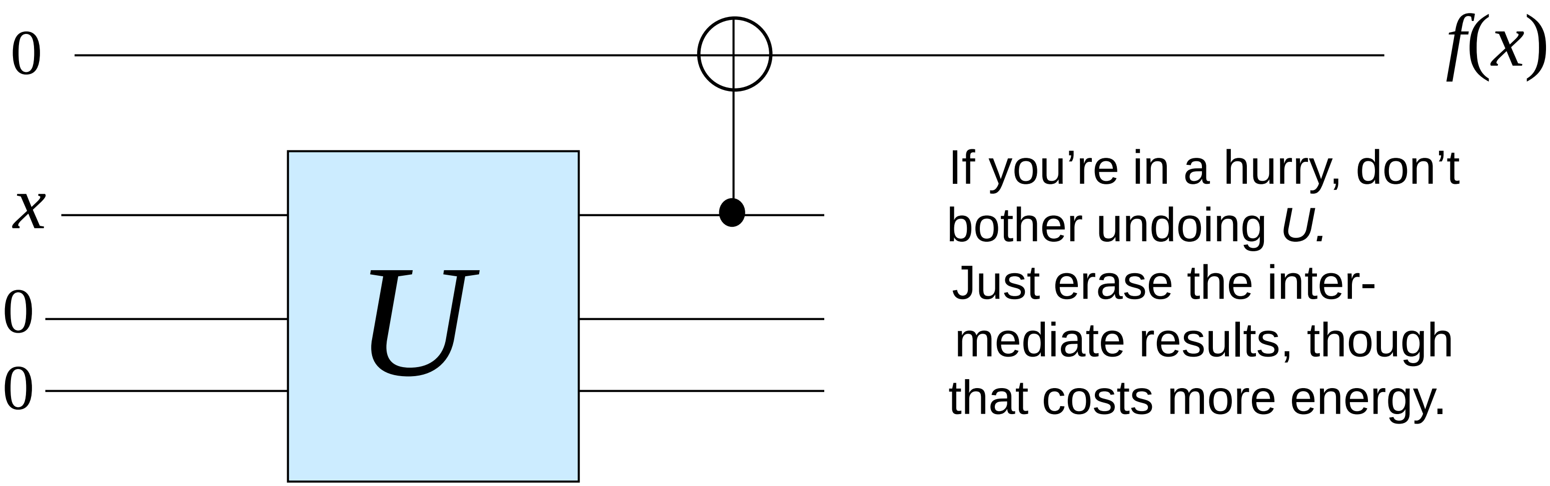
Conventional　Efficient 1:1 map　Efficient 2:1 map



Cost E_B >>kT　Cost 0　Cost kT ln 2

Why should there be a need, or desire, to erase garbage bits?

After all, the theory of reversible computation shows that deterministic computations, if they're allowed to save a copy of the input, need produce no garbage at all, since all unwanted intermediate data can be disposed of by undoing the process that created it. Similarly any quantum computation of a deterministic function can be embedded in a two-stage unitary computation that regenerates the input and produces no garbage.



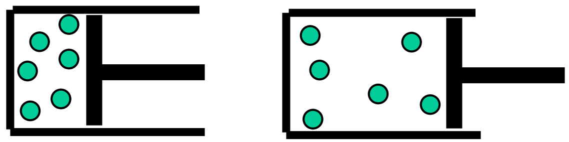
If you're in a hurry, don't bother undoing *U*. Just erase the intermediate results, though that costs more energy.

Answer: logically irreversible erasure is typically used as a quick and dirty way of getting rid of unwanted but *determinate* data. It's like throwing away stuff that with more patience could have been recycled.

Three Kinds of Entropy are more similar than they first appear

- **Thermodynamic** entropy difference between equilibrium states

$$\Delta S = \int_{\text{rev.}} dQ/T$$



Relation to statistics: isothermally compressing a gas to half its volume transfers N bits of entropy to environment, with the reversible flow of NkT log 2 of heat.

- **Statistical** entropy of a distribution $\boldsymbol{P}$ or mixed state $\boldsymbol{\rho}$

$H(\boldsymbol{P}) = -\Sigma_x P(x) \log P(x)$, or quantumly $H(\boldsymbol{\rho}) = -\mathrm{Tr}\, \boldsymbol{\rho} \log \boldsymbol{\rho}$

- **Algorithmic** entropy or Kolmogorov complexity of a bit string x

$K(x) = \min \{|p|: U(p)=x\}$ i.e. the size in bits of the smallest program p causing a standard universal computer U to compute exactly x as output.

If $\boldsymbol{P}$ is a concisely describable ensemble, its *statistical* entropy is nearly equal to the average *algorithmic* entropy of its members.

$H(\boldsymbol{P}) < \Sigma_x P(x) K(x) < H(\boldsymbol{P}) + K(\boldsymbol{P})$. (Zvonkin & Levin '70, cf also Bennett '82)

The integration of physical, especially quantum, ideas and methods has been fairly successful in informatics.

But 21st century **cosmology** offers new challenges and opportunities for intercultural sensitivity and synthesis. Observational astronomy strongly supports the ΛCDM "standard model", which predicts that the expansion of our universe is accelerating, leading to an infinite future at thermal equilibrium at a positive but very low temperature. This so-called asymptotic de Sitter state raises two fundamental questions:

• The Boltzmann brain problem—how do we know we are inhabitants of a young live universe rather than fluctuations in an old dead one?

• The Wigner's friend problem—what does it feel like to be inside a quantum superposition? In particular, does the de Sitter state even have fluctuations, if there is no measuring apparatus present to observe them?

- We have already reviewed the relation between *Dynamics*—the spontaneous motion or change of a system obeying physical laws—and *Computation*—a programmed sequence of mathematical operations

- Self-organization, exemplified by cellular automata and *logical depth* as a measure of complexity.

- True and False evidence—the Boltzmann Brain problem at equilibrium and in modern cosmology

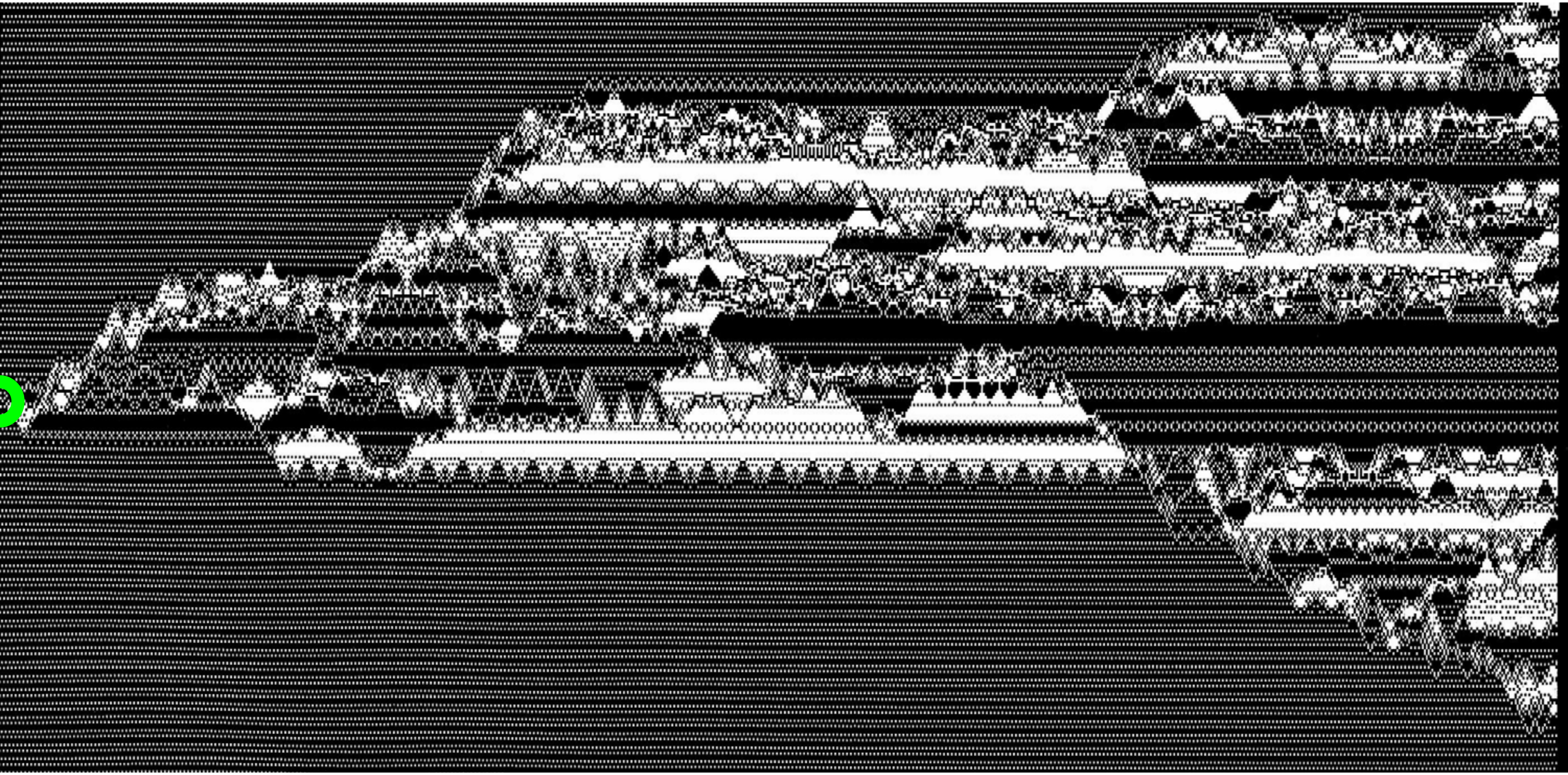
- Wigner's Friend—what it feels like to be inside an unmeasured quantum superposition

How does the familiar complicated world we inhabit emerge *cosmologically* from the austere high-level laws of quantum mechanics and general relativity, or *terrestrially* from lower-level laws of physics and chemistry?

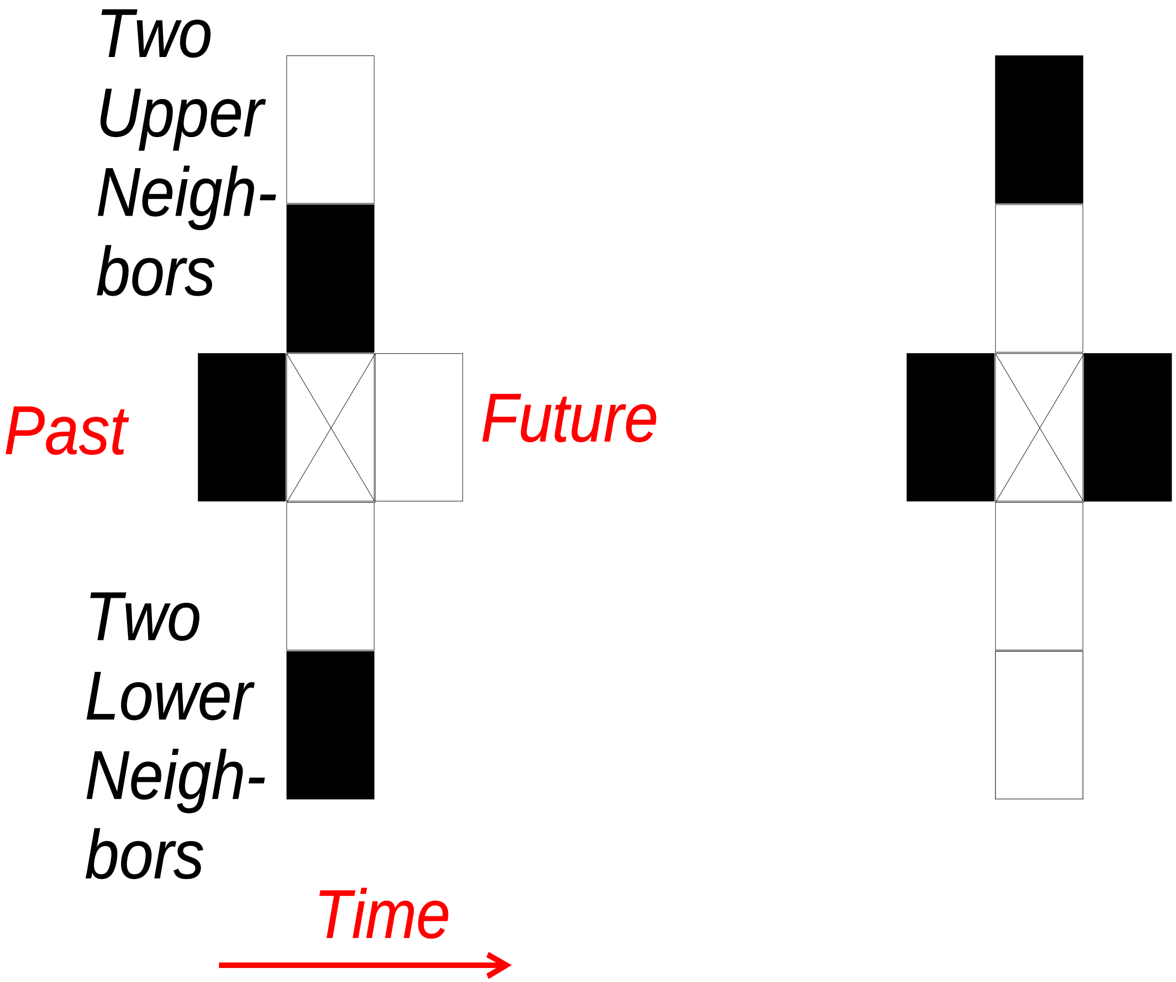
To attack this question in a disciplined fashion, one must first define complexity, the property that increases when a self-organizing system organizes itself.

(I am fairly new to cosmology, and would welcome advice from experts in case some of the questions I ask are ill-posed, or the answers already known.)

Simple classical dynamics (such as this 1 dimensional reversible cellular automaton) are easier to analyze and can produce structures of growing “complexity” from simple initial conditions. time ⟶

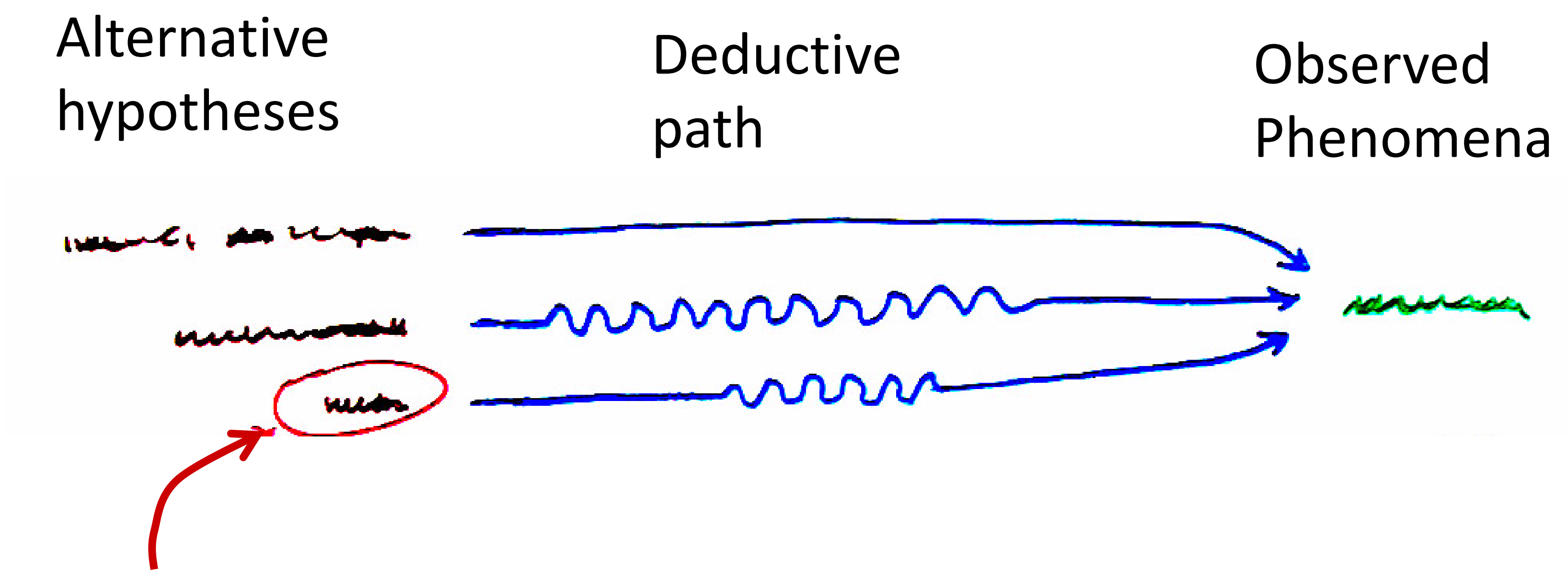


Small irregularity (green) in otherwise periodic initial condition produces a complex deterministic wake.



Range-2, deterministic, 1-dimensional Ising rule. Future differs from past if exactly two of the four nearest upper and lower neighbors are black and two are white at the present time.

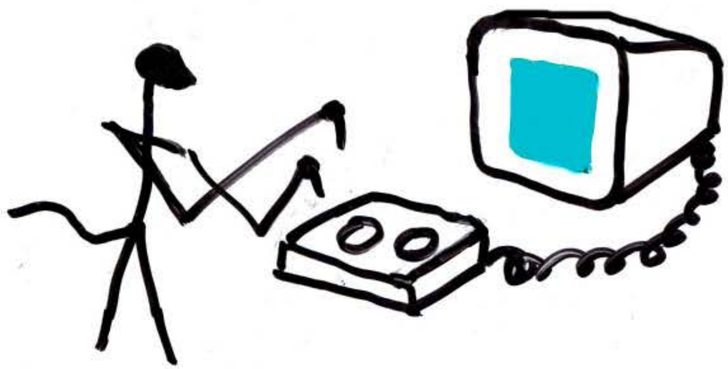
Occam's Razor



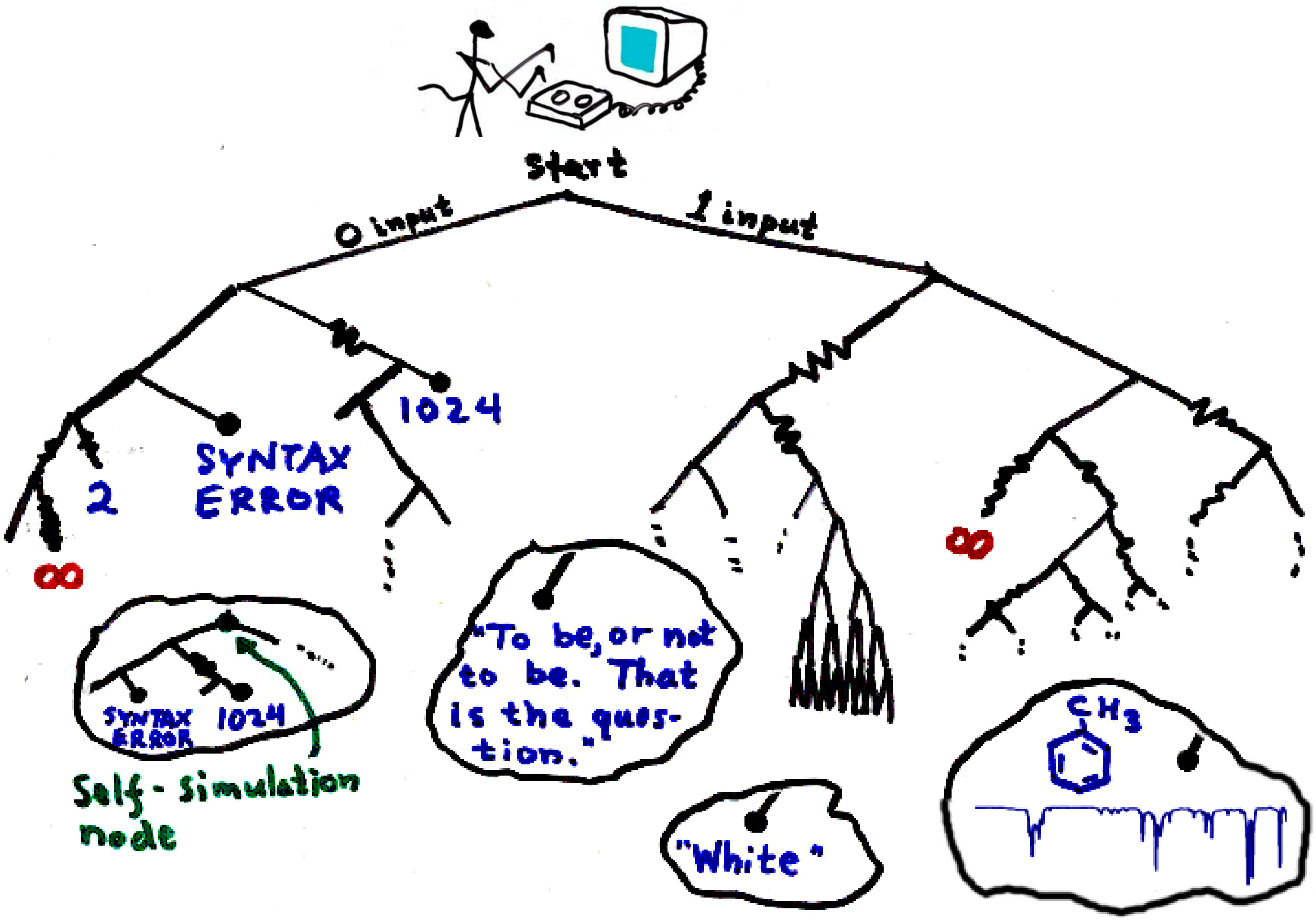
The most economical hypothesis is to be preferred, even if the deductive path connecting it to the phenomena it explains is long and complicated.

But how does one compare economy of hypotheses in a disinterested way?

Algorithmic information, devised in the 1960's by Solomonoff, Kolmogorov, and Chaitin, uses a computerized version of the old idea of a monkey at a typewriter eventually typing the works of Shakespeare.

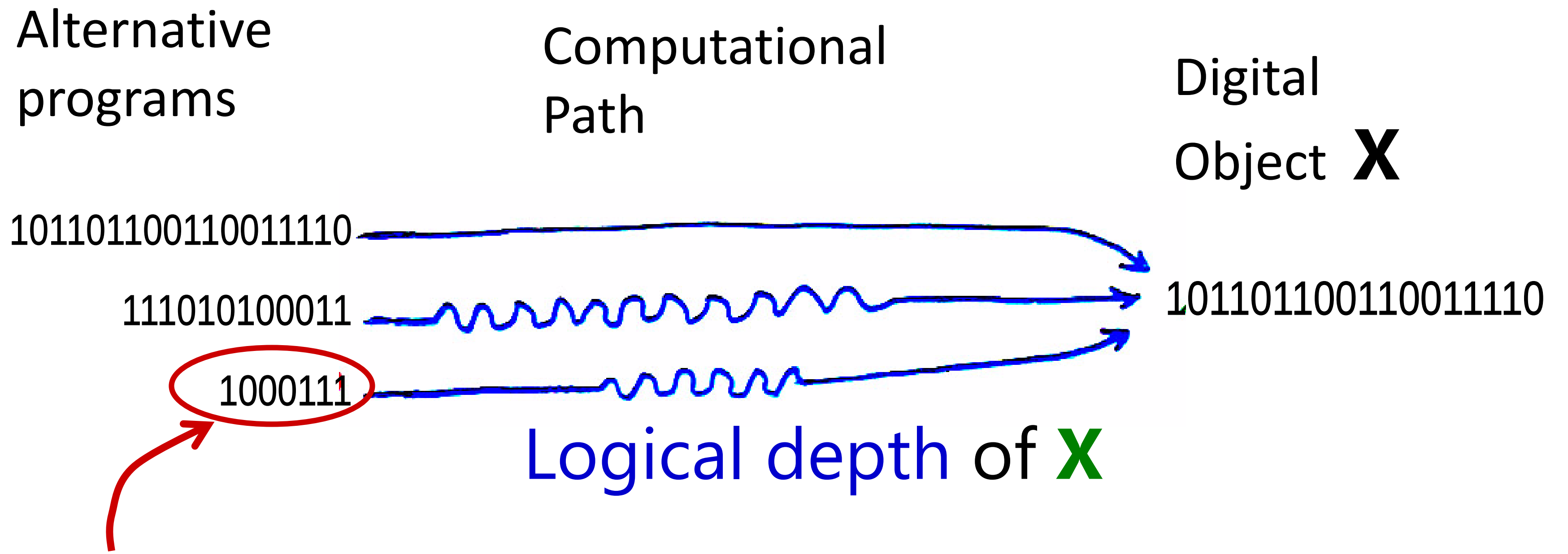


A monkey randomly typing 0s and 1s into a universal binary computer has some chance of getting it to do any computation, produce any output.



This tree of all possible computations is a microcosm of all cause/effect relations that can be demonstrated by deductive reasoning or numerical simulation.

In a computerized version of Occam's Razor, the hypotheses are replaced by alternative programs for a universal computer to compute a particular digital (or digitized) object **X**.



The shortest program is most plausible, so its *run time* measures the object's logical depth, or plausible amount of computational work required to create the object.

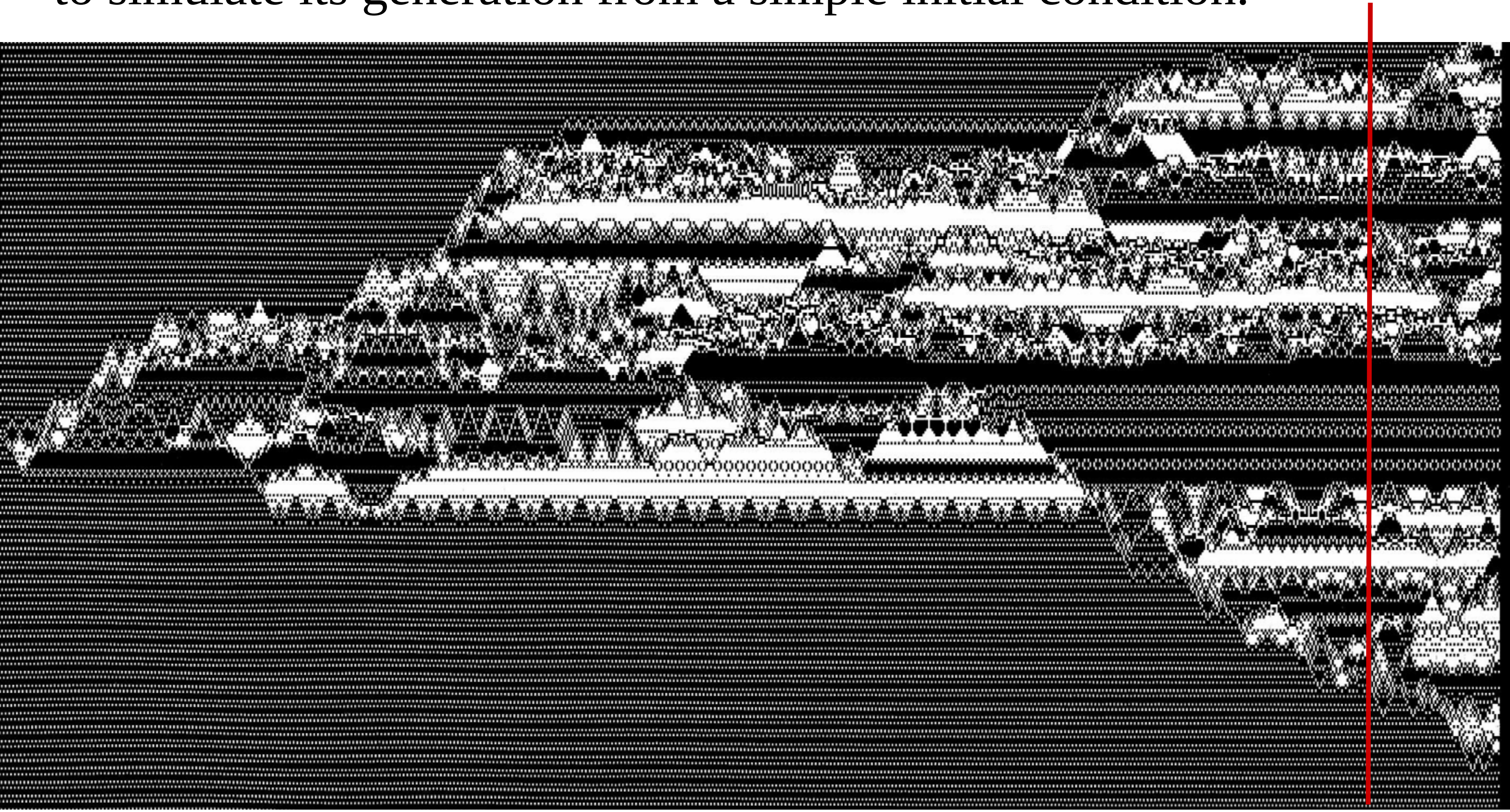
A trivially orderly sequence like 111111… is logically shallow because it can be computed rapidly from a short description.

A typical random sequence, produced by coin tossing, is also logically shallow, because it essentially **its own** shortest description, and is rapidly computable from that.

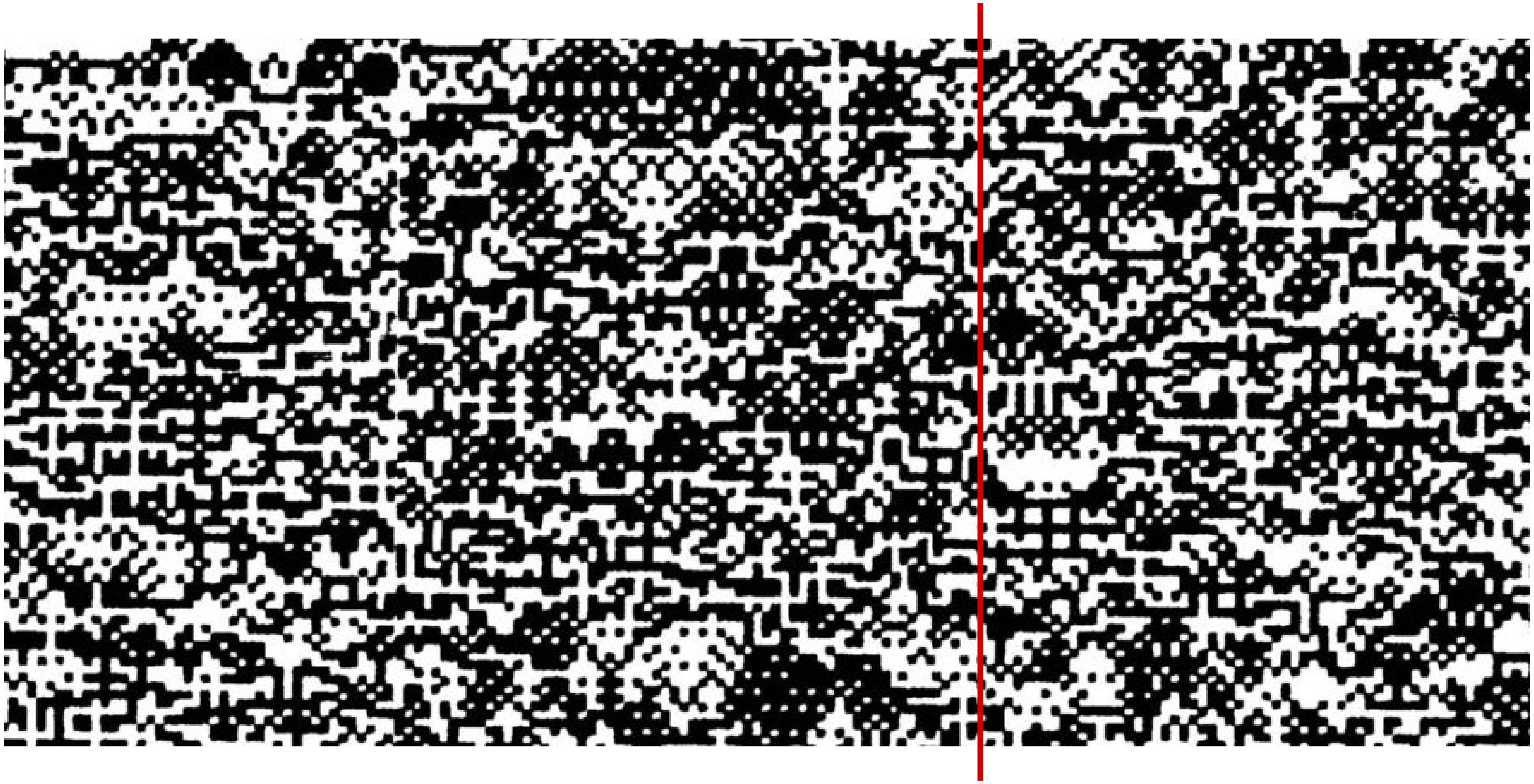
Trivial semi-orderly sequences, such as an alternating sequence of 0's and random bits, are also shallow, since they are rapidly computable from their random part.

(Depth is thus distinct from, and can vary independently from *Kolmogorov complexity* or *algorithmic information content*, defined as the **size** of the minimal description, which is high for random sequences. Algorithmic information measures a sequence's randomness, not its complexity in the sense intended here.)

Initially, and continuing for some time, the logical depth of a time slice increases with time, corresponding to the duration of the slice's actual history, in other words the computing time required to simulate its generation from a simple initial condition.



But if the dynamics is allowed to run for a large random time after equilibration (comparable to the system's Poincaré recurrence time, exponential in its size), the typical time slice becomes shallow and random, with only short-range correlations.



The minimal program for this time slice does not work by retracing its actual long history, but rather a short computation short-circuiting it.

Why is the true history no longer plausible?

Because to specify the state via a simulation of its actual history would involve naming the exact **number** of steps to run the simulation.

This number is typically very large, requiring about $\boldsymbol{n}$ bits to describe.

Therefore the actual history is no more plausible (in terms of Occam’s razor) than a “print program” that simply outputs the state from a verbatim description.

In a world at thermal equilibrium, with local interactions, correlations are generically local, mediated through the present.

Correlations mediated through present only

By contrast, in a non-equilibrium world, local dynamics can generically give rise to long range correlations, mediated through a V-shaped path in space-time representing a common history.

time

Elizabeth I

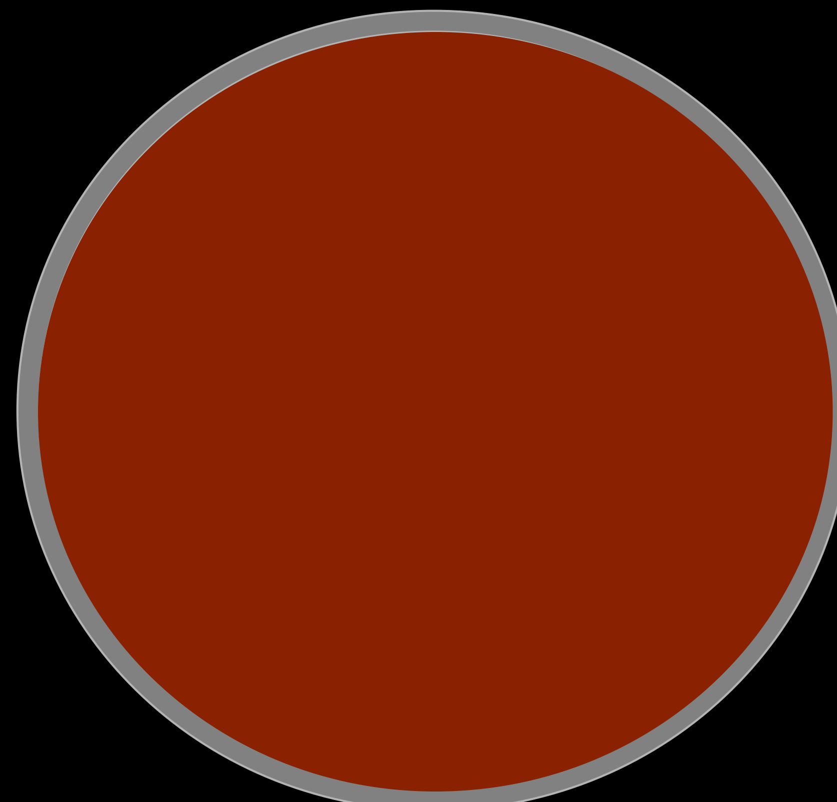
Grenada
1999

Elizabeth II

Canada
2002

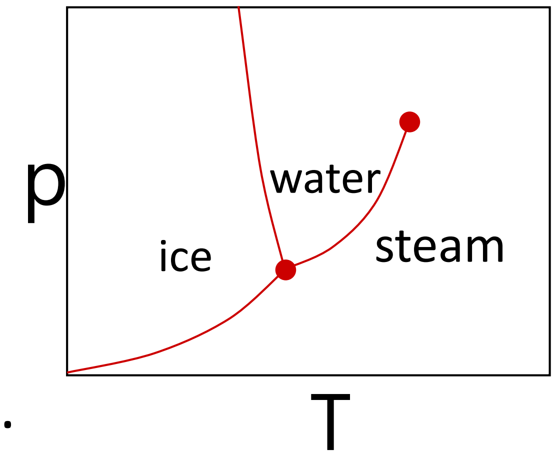
The cellular automaton is a classical toy model, but real systems with fully quantum dynamics behave similarly, losing their complexity, their long-range correlations and even their classical phenomenology as they approach equilibrium.

If the Earth were put in a large reflective box and allowed to come to equilibrium, its state would no longer be complex or even phenomenologically classical.

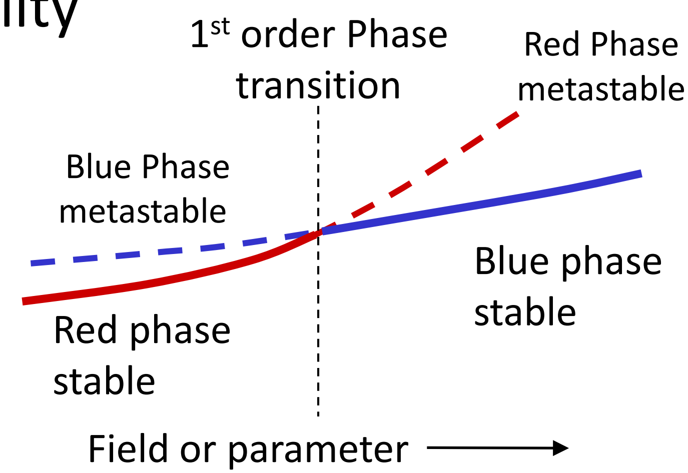


The entire state in the box would be a microcanonical superposition of near-degenerate energy eigenststates of the closed system. Such states are typically highly entangled and contain only short-range correlations.

How strong is the connection between disequilibrium and complexity, in the sense of logical depth?



Are thermal equilibrium states generically shallow? Classically Yes, by the Gibbs phase rule. For generic parameter values, a locally interacting classical system, of finite spatial dimensionality and at finite temperature, relaxes to a unique phase of lowest bulk free energy.

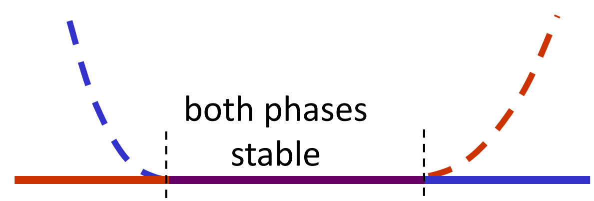


=> no long term memory

=> depth remains bounded in large N limit

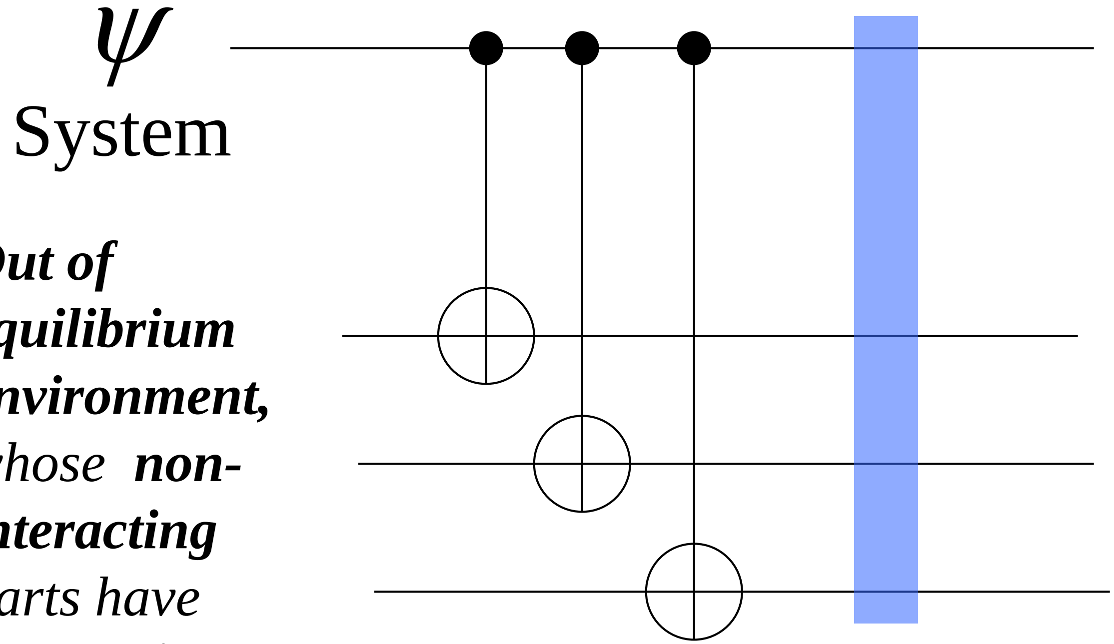
Quantum Exceptions? Toric code in 3 or more dimensions, Localization

Dissipative systems are exempt from the Gibbs phase rule (BG '85)



d+1 dimensional "free energy" from system's transfer matrix

How eavesdropping by the environment creates a classical-appearing world.



Out of equilibrium environment, *whose* ***non-interacting*** *parts have* ***commuting*** *interactions with the system.*

Massive eavesdropping causes the system to get ***classically correlated*** *with may* ***individual parts*** *of its environment. But because of the monogamy of entanglement, it remains entangled only with the* ***whole environment.***

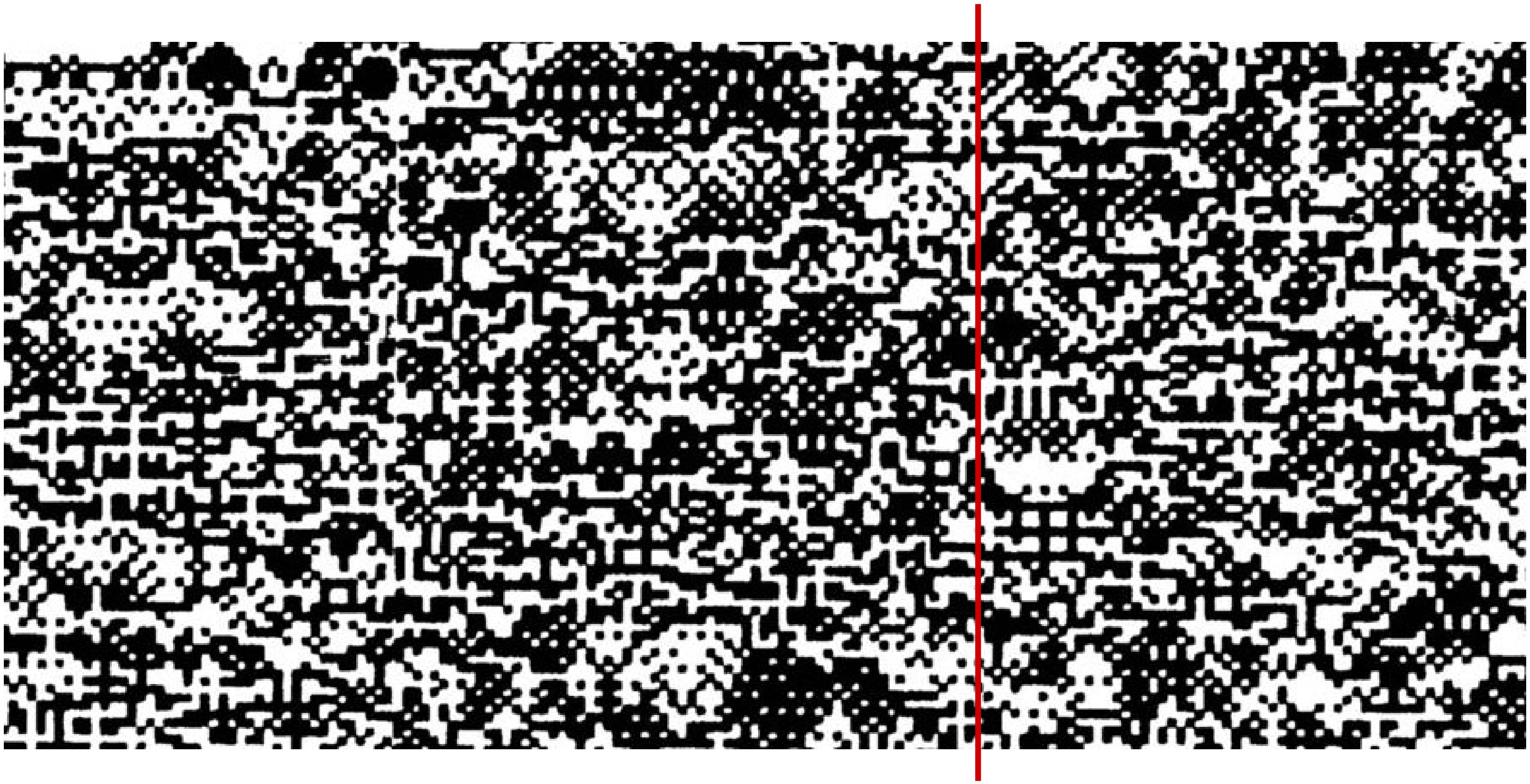
This brand of decoherence theory is called "Quantum Darwinism" (Zurek et al), but a better name would be Quantum Spam, since the multiple copies all come from a single original.

Riedel and Zurek have pointed out the role of non-thermal illumination in creating classical correlations in everyday life, e.g. photons from the sun reflecting off objects on the surface of the Earth to produce massively redundant records of their positions.

If these photons continue to propagate away in free space, the system will never equilibrate and the redundant records will be permanent, though inaccessible, even outliving the Earth.

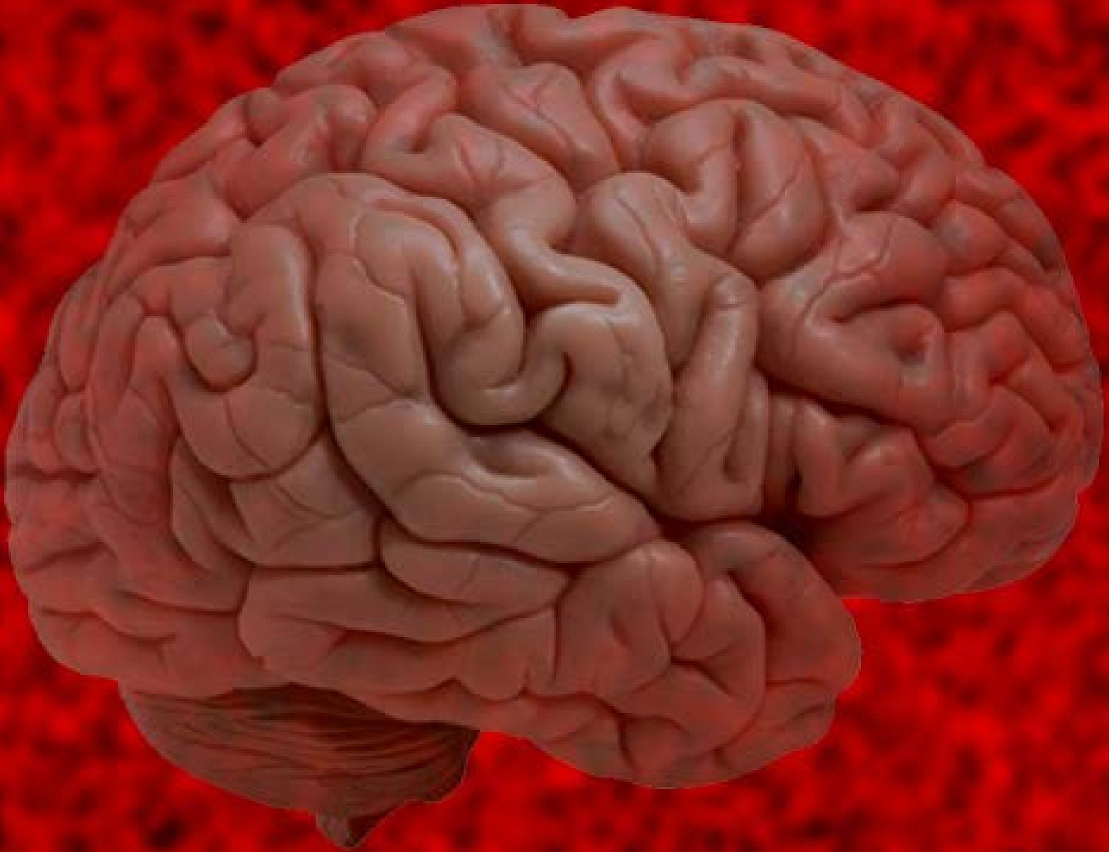
But if the reflected photons were instead trapped inside a reflective box, they would be repeatedly absorbed and re-emitted from the Earth, obfuscating the former redundant correlations as the system equilibrates, and rendering the system no longer classical.

Recall that if a system's dynamics is allowed to run for a long time after equilibration (comparable to the system's Poincaré recurrence time) its actual history can no longer be reliably inferred from its present state.



Conversely, a deep structure, one that seems to have had a long history, might just be the result of an unlikely thermal fluctuation, a so-called Boltzmann Brain.

A friend of Boltzmann proposed that the low-entropy world we see may be merely a thermal fluctuation in a much larger universe. “Boltzmann Brain” has come to mean a fluctuation just large enough to produce a momentarily functioning human brain, complete with false memories of a past that didn’t happen, and perceptions of an outside world that doesn’t exist. Soon the BB itself will cease to exist.



Nowadays serious cosmologists worry about Boltzmann Brains

e.g. arxiv:1308.4686

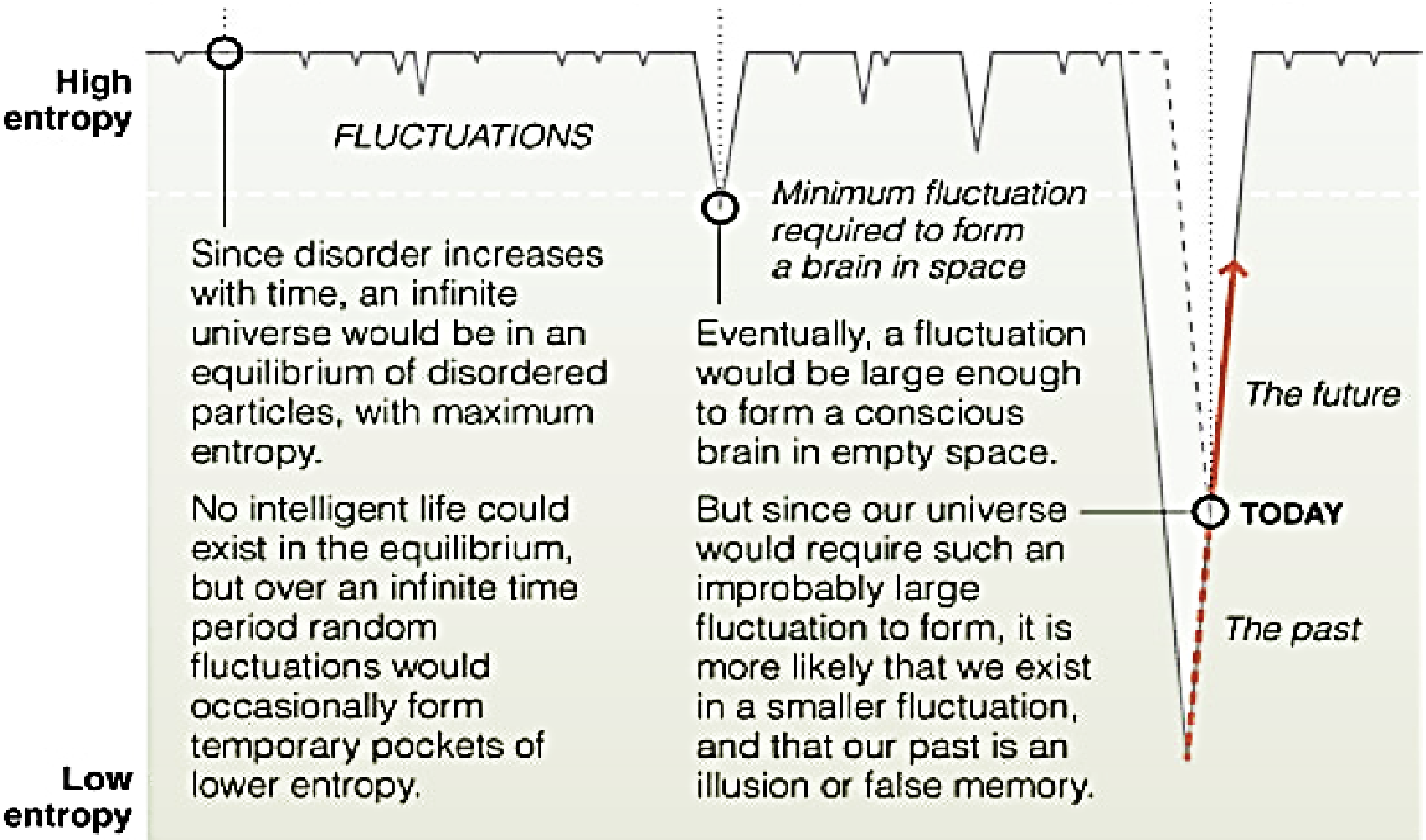
Can the Higgs Boson Save Us From the Menace of the Boltzmann Brains?

Kimberly K. Boddy and Sean M. Carroll

California Institute of Technology, Pasadena, CA 91125

The standard ΛCDM model provides an excellent fit to current cosmological observations but suffers from a potentially serious Boltzmann Brain problem. If the universe enters a de Sitter vacuum phase that is truly eternal, there will be a finite temperature in empty space and corresponding thermal fluctuations. Among these fluctuations will be intelligent observers, as well as configurations that reproduce any local region of the current universe to arbitrary precision. We discuss the possibility that the escape from this unacceptable situation may be found in known physics: vacuum instability induced by the Higgs field. Avoiding Boltzmann Brains in a measure-independent way requires a decay timescale of order the current age of the universe, which can be achieved if the top quark pole mass is approximately 178 GeV. Otherwise we must invoke new physics or a particular cosmological measure before we can consider ΛCDM to be an empirical success.

A diabolical conundrum: Boltzmann fluctuations nicely explain the low entropy state of our world, and the arrow of time, but they undermine the scientific method by implying that our picture of the universe, based on observation and reason, is ***false.***



Source: Sean Carroll, California Institute of Technology

JONATHAN CORUM/THE NEW YORK TIMES

Diabolical Conundrum Continued: People began worrying about equilibration in the 19th Century, calling it the "heat death of the universe", but thought of it as a problem for the far future.

Boltzmann showed us that it is already a problem in the present, undermining our ability to make inferences about conditions in the past or elsewhere, based on those here and now. The inhabitants of any universe that will ultimately equilibrate, either microcanonically or canonically, must make the additional postulate, unsupported by observation, that they are situated **atypically early** in its history. Otherwise, their "scientific" inferences are no better than those of the inhabitants of Borges' fictional *Library of Babel* (which contained, randomly shelved, one copy of each possible 410 page book).

Wigner’s Friend

Eugene Wigner imagined a gentler version of Schrödinger’s Cat, relevant to the Quantum Boltzmann Brain problem:

Wigner’s friend performs a quantum measurement with two outcomes but only tells Wigner what happened later.

After the experiment, but before Wigner hears the result, Wigner regards his friend as being in a superposition of two states, but the friend perceives only one or the other of them.

In principle (and even in practice, for atom-sized friends) Wigner can contrive for the friend to undo the measurement and forget its result—a “quantum eraser” experiment.

Wigner's friend might have been viewed as no more than a philosophical conundrum, but it is relevant to the anthropic counting of observers.

In a 2014 sequel to their 2013 paper, Boddy and Carroll, joined by Pollack, argue that it is not necessary for the universe to self-destruct to avoid the menace of Boltzmann brains. They instead argue that the late thermal state of the universe doesn't generate any Boltzmann brains because there is no mechanism to **observe** them, in the strong sense of making a permanent external classical record.

But as Jess Riedel and I have argued, all our experience, like that of Wigner's friend, is potentially impermanent. Therefore I think it is unreasonable to insist that nothing happens until a permanent record of it is made. Moreover observership, in the anthropic sense, is an introspective property of a system, not a property of how it would behave if measured externally.

If a piece of our universe, centered on the sun, were put in a box with perfectly reflective walls, 1 million light years in diameter, it would take us half a million years to notice any difference. Yet the long term evolution of this isolated system would be radically different from the evolution of the universe we believe we inhabit, lacking this box. The boxed universe would recur repeatedly to near its initial state, and, exponentially more frequently, to Boltzmann brain states, where the recurrence would be confined to a solar-system sized patch, or smaller. So unless one is willing to push the moveable quantum-classical boundary out indefinitely far, this system would experience what we experience now, but on its orbit false local recurrences would vastly outnumber true ones.

Similarly, we argue, in the thermal de Sitter state of an unboxed universe, false local recurrences would vastly outnumber full recurrences, and these would infinitely outnumber the single first-time occurrence of our solar system in the young expanding universe.

To think about this, it helps to review some basic facts about entanglement and quantum mixed states:

•A mixed state is completely characterized by its density operator ρ, which describes all that can be learned by measuring arbitrarily many specimens of the state. For an ensemble of pure states, $\{p_j, \psi_j\}$, ρ is given by the weighted sum of the projectors onto these states.

•Ensembles with the same ρ are indistinguishable.

•A system **S** in a mixed state $\rho^{\mathbf{S}}$ can, without loss of generality, be regarded as a subsystem of a larger bipartite system **RS** in a pure state $\Psi^{\mathbf{RS}}$, where R denotes a non-interacting reference system.

•"Steering" Any ensemble $\{p_j, \psi_j\}$ compatible with ρ can be remotely generated by performing measurements on the R part of $\Psi^{\mathbf{RS}}$. Measurement outcome j occurs with probability p_j, leaving S in state ψ_j.

Jess Riedel's scenario suggesting why Boltzmann brains ought to be present in thermal states at any positive temperature, even though there is no external observer.

- Let π_{BB} be a projector onto some state representing a fluctuation, for example a copy of the Solar System pasted into a much larger patch of de Sitter vacuum.
- Any finite temperature thermal state $\boldsymbol{\rho}$ of this patch can be expressed as a weighted sum

$$\rho = \lambda\, \pi_{BB} + (1-\lambda)\, \boldsymbol{\sigma}$$

where $\boldsymbol{\sigma}$ is a thermal state "depleted" in $\boldsymbol{\pi}_{BB}$.

- An all-powerful Preparator tosses a λ-biased coin, and prepares π_{BB} or $\boldsymbol{\sigma}$ according to the outcome.
- Before departing, the Preparator takes away, in reference system **R**, a record of all this, including, for example, souvenir photos of the just-created Earth and its inhabitants.

Since this is a valid preparation of the thermal state, and keeping in mind that it is impossible in principle to distinguish different preparations of the same mixed state, it is hard to see why the inhabitants of the de Sitter patch do not have some small probability of experiencing a life resembling our own, at least for a while.

Jason Pollack's reply to this argument: their 2014 paper, alleging the absence of such fluctuations, does not apply to all thermal states, but only those purified by a reference system **R** of a particular form, so that state $\Psi^{\mathbf{RS}}$ is a Bunch-Davies pure state of the universe whose local patches $\rho^{\mathbf{S}}$ are all in thermal de Sitter states.

This may be viewed as an Occam-type argument from simplicity, favoring simplicity not of the accessible system **S**, but of the largely inaccessible **RS**.

Internal vs External views: Our suggested internal criterion for a state $\boldsymbol{\rho}$ to have nonzero participation of a Boltzmann brain state π_{BB}, namely

$$\exists \boldsymbol{\sigma}, \lambda > 0: \quad \rho = \lambda\, \pi_{BB} + (1-\lambda)\, \boldsymbol{\sigma}$$

is more restrictive than the usual criterion that $\boldsymbol{\rho}$ have a positive expectation when subjected to an external measurement of π_{BB}, namely,

$$\mathrm{tr}(\rho\, \pi_{BB}) > 0.$$

Even a zero temperature vacuum state (the Lorentz vacuum) would have a positive Boltzmann brain probability when measured externally. The energy for creating the Boltzmann brain out of the ground state would come from the measuring apparatus. This is a further reason we think an external measuring apparatus is an encumbrance in a cosmological setting, when reasoning about a system's internal experiences.

Open questions

- Wigner's Friend's experiences, if any, and their relevance to the counting of observers
- Do entanglement and topological order enable generic fault-tolerant memory and self-organization at equilibrium (escape from Gibbs phase law)
- If not, are there cosmologies (e.g. eternal inflation) providing perpetual disequilibrium sufficient to support unbounded fault-tolerant classical self-organization

Einstein’s greatest achievement, general relativity, allows the existence of severely warped spacetimes containing closed timelike curves (CTCs). If such curves are sufficiently stable, they might make some form of time travel possible. In another cultural interaction with physicists, some computer scientists have suggested that equipping a quantum computer with a CTC would dramatically enhance its computational and state-discrimination powers.

Leung, Smith, Smolin and I have disputed these conclusions, arguing that the are based on inconsistent ways of formalizing the notion of a computational task (arXiv:0908.3023)

Workshop on “Quantum Foundations of a Classical Universe,” IBM Research Aug 11-14, 2014
http://www.jessriedel.com/conf2014/conf2014.html or
http://researcher.watson.ibm.com/researcher/view_group.php?id=5661

C. J. Riedel and W. H. Zurek, "Quantum Darwinism in an Everyday Environment: Huge Redundancy in Scattered Photons," *Phys. Rev. Lett.* **105**, 020404 (2010). [arXiv:1001.3419] cf also longer treatment in [arxiv:1102.31793v3]

C.J. Riedel, Classical branch structure from spatial redundancy in a many-body wavefunction, arXiv:1608.05377.

C.H. Bennett blog post on logical depth versus other complexity measures http://dabacon.org/pontiff/?p=5912

CH Bennett, blog post on *Schopenhauer and the Geometry of Evil, https://quantumfrontiers.com/2016/05/29/schopenhauer-and-the-geometry-of-evil/*

C.H. Bennett "Logical Depth and Physical Complexity" in *The Universal Turing Machine– a Half-Century Survey*, edited by Rolf Herken Oxford University Press 227-257, (1988)
http://researcher.ibm.com/researcher/files/us-bennetc/UTMX.pdf

C.H. Bennett and G. Grinstein "On the Role of Dissipation in Stabilizing Complex and Non-ergodic Behavior in Locally Interacting Discrete Systems" *Phys. Rev. Lett.* **55,** 657-660 (1985).
http://researcher.ibm.com/researcher/files/us-bennetc/BG85%20with%20Toom%20snapshotsq.pdf

Peter Gacs, “Reliable Computation with Cellular Automata” *J. Computer and System Science* **32**, 15-78 (1986)
http://www.cs.bu.edu/~gacs/papers/GacsReliableCA86.pdf
Bennett, Leung, Smith & Smolin arXiv:0908.3023 and PRL, (Critique of alleged computational speedups from Closed timelike curves.)

*Answers to questions about Gabor's high-compression Szilard engine.

Correct answer: No trapping mechanism, whether mechanical (e.g. a mouse trap) or optical (Gabor's engine), can be completely irreversible. By the principle of Smoluchowski's trap door and Feynman's ratchet, the work $\mathbf{w}$ of resetting a trap, rather than being constant, must increase logarithmically with the compression ratio V/V_0, to keep the trap from running in the reverse of its intended direction.

Gabor's 1961 answer instead invoked quantum optics, saying the longer the cylinder the more optical modes it has, and the more energy would be required to confine a light beam to one end of it. Though true, this implied that that quantum effects were necessary to save the second law, whereas simple considerations of reversibility suffice.