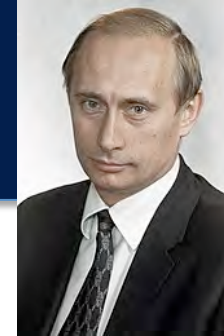# Living Digital Beings

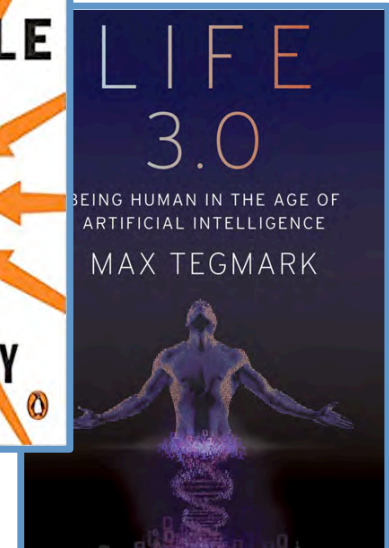*Edward A. Lee*

UC Berkeley

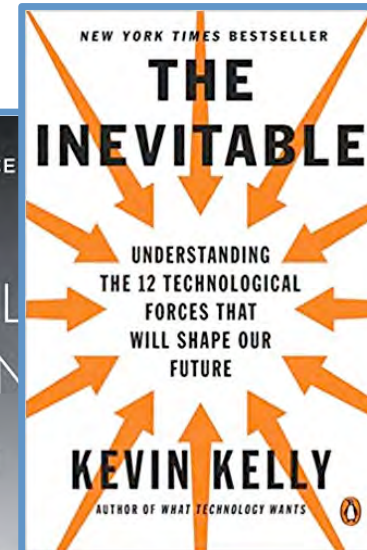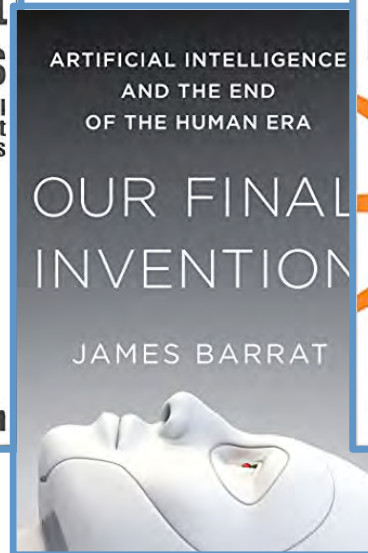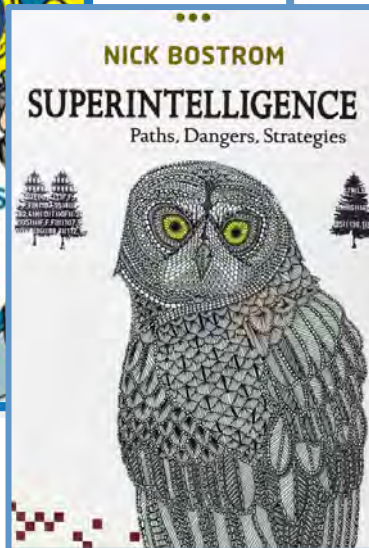*Invited talk, VERIMAG, April 1, 2019*

University of California at Berkeley

Hype and Fear

Is AI an *existential threat* to humanity?

Three Questions About AIs

1. Are we going to lose control of them?

2. Are they alive?

3. Are they going to match and exceed us?

Changing the Question

Can we teach computers to program?

–Maarten van Steen, March 19, 2019

Can computers teach humans to program?

–Edward Lee, March 20, 2019

Computers Teaching Humans to Program

Eclipse

Jupyter

Github

Stack Overflow

Google

...

Eggs and Chickens

Richard Dawkins

"A chicken is an egg's way of making another egg."

Is a human a computer's way of making another computer?
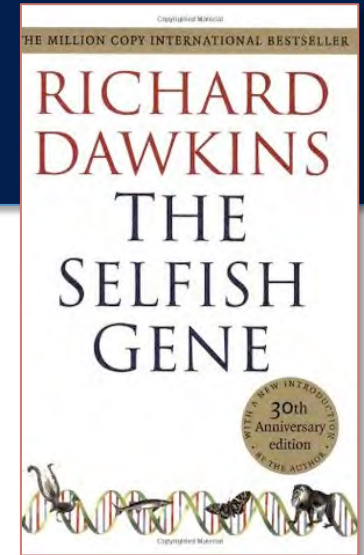
Digital Creationism: The Hypothesis that Technology is Top-Down Intelligent Design

Vasa

Evolutionary processes are capable of much more complex and sophisticated design than top-down intelligent design.
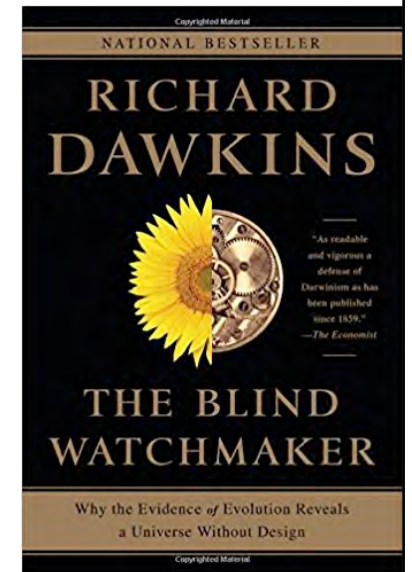
An Alternative to Digital Creationism: Symbiotic Coevolution

The Creative Partnership of Humans and Technology

PLATO AND THE NERD

EDWARD ASHFORD LEE

"Are we playing God, creating a new life form in our own image, or are we being played by a Darwinian evolution of a symbiotic new species?"

"Are humans the purveyors of the 'noisy channel' of mutation, facilitating sex between software beings by recombining and mutating programs into new ones?"

8
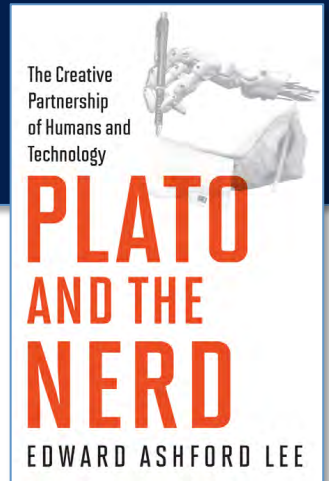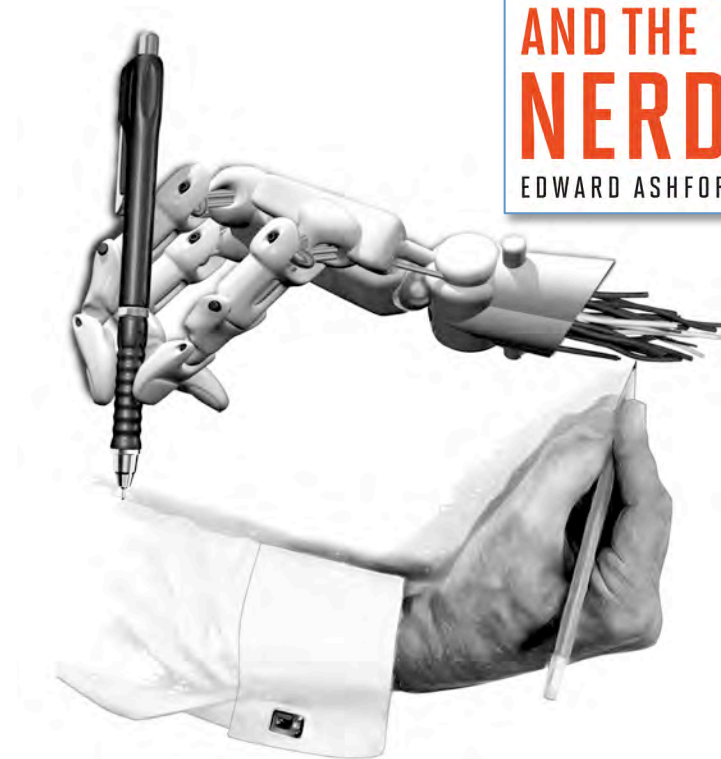
Coevolution of Digital Beings and Humans

[Photo By James Morrison CC BY-SA 2.0]

DARWIN among the MACHINES
The Evolution of Global Intelligence

"As remarkable an intellectual history as any I have read."—OLIVER SACKS

George Dyson

1997

Turing's Cathedral

THE ORIGINS OF THE DIGITAL UNIVERSE

George Dyson

2012

9
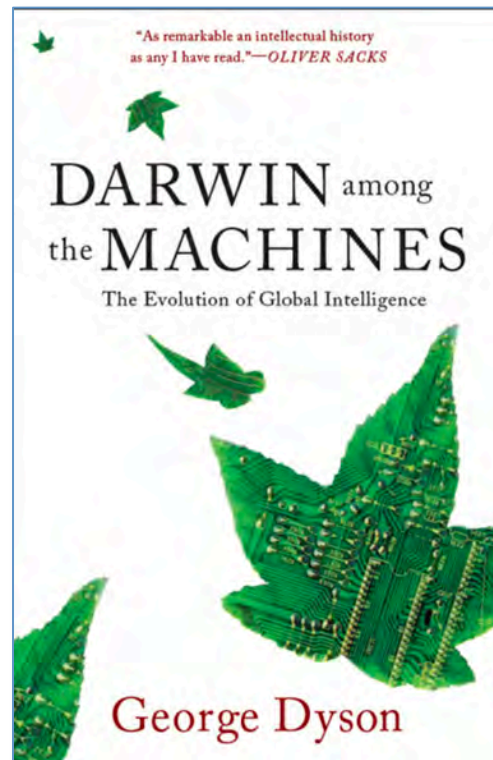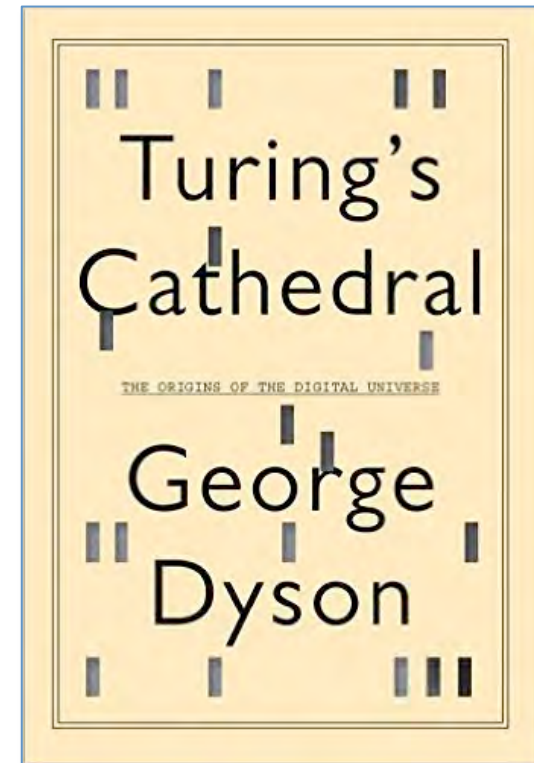
Humans in Control of AI?

Are we going to lose control of them?

No.

We never were in control, so we can't lose control.

Three Questions

1. Are we going to lose control of them?
2. Are they alive?
3. Are they going to match and exceed us?

The Technium

Kevin Kelly, talks about the "technium" as the 7th kingdom of life.
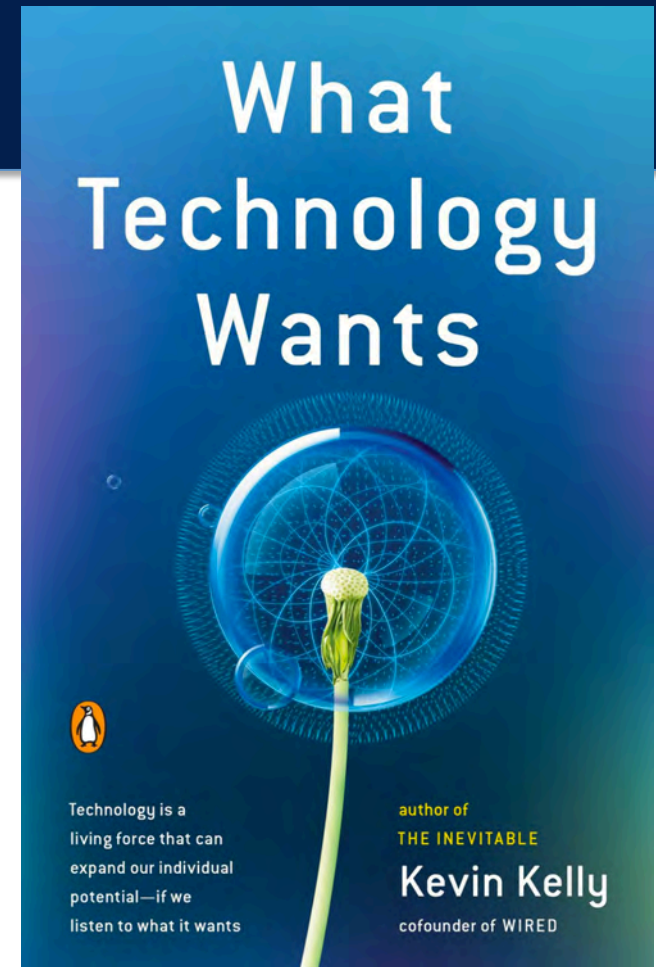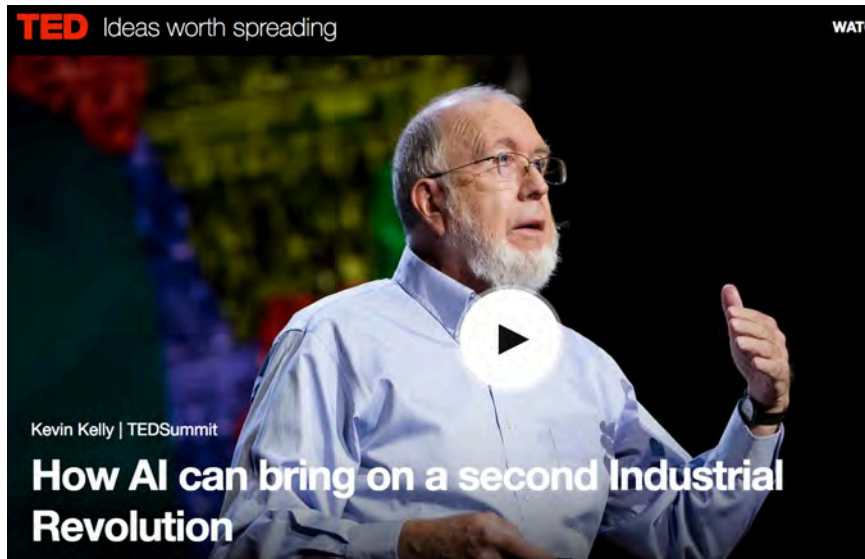
2010

Is It Alive?
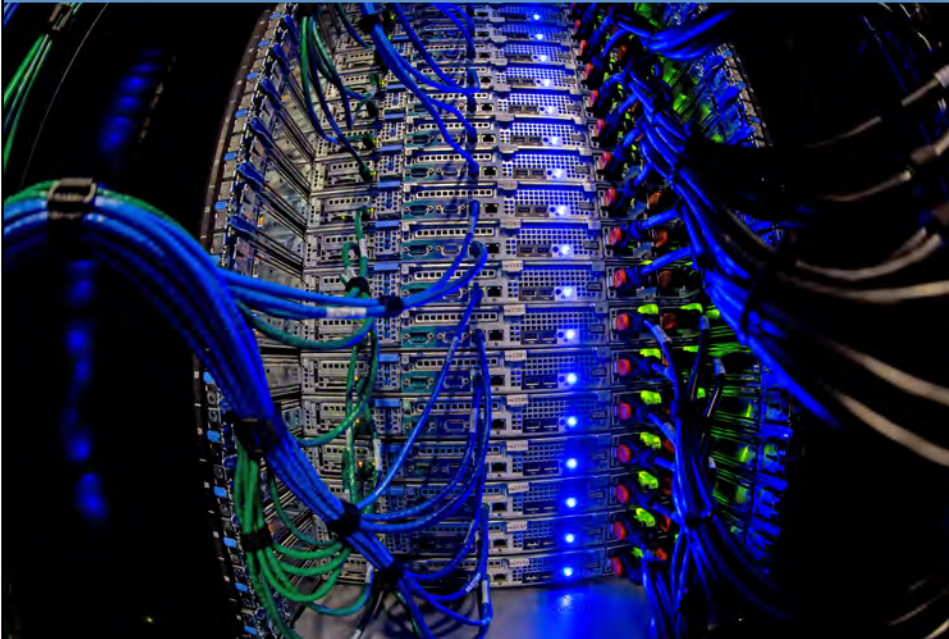
Wikipedia Servers
[Victor Grigas/Wikimedia Foundation CC BY-SA 3.0]
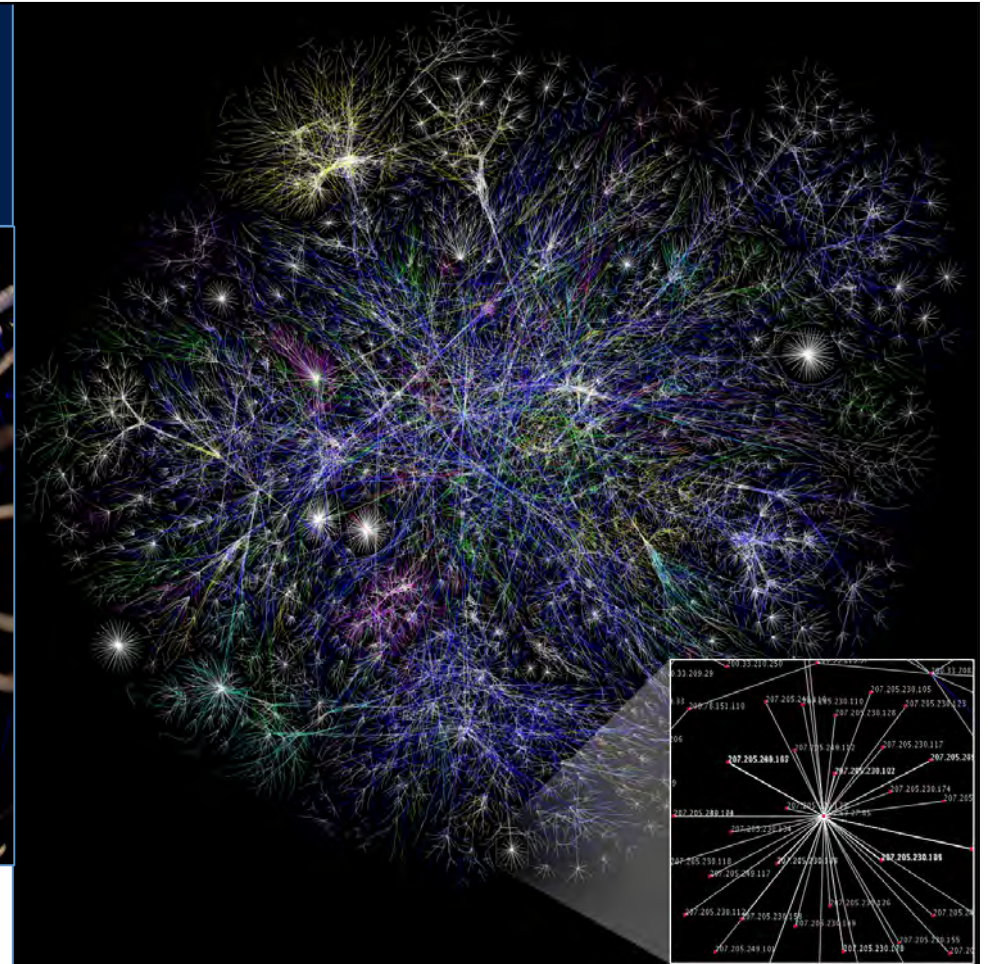
View of the Internet
[The Opte Project, via Wikimedia Commons CC BY 2.5]

13

What does it mean to be alive?

Wikipedia is arguably a "living digital being" (LDB, or "eldebee").

It has all of these properties.

GROWTH
and development

RESPOND
to the environment

REPRODUCE
offspring

ALIVE

CELLULAR
structure & composition

HEREDITY
of traits

METABOLISM
energy from
chemical reactions

HOMEOSTASIS
stable inner conditions
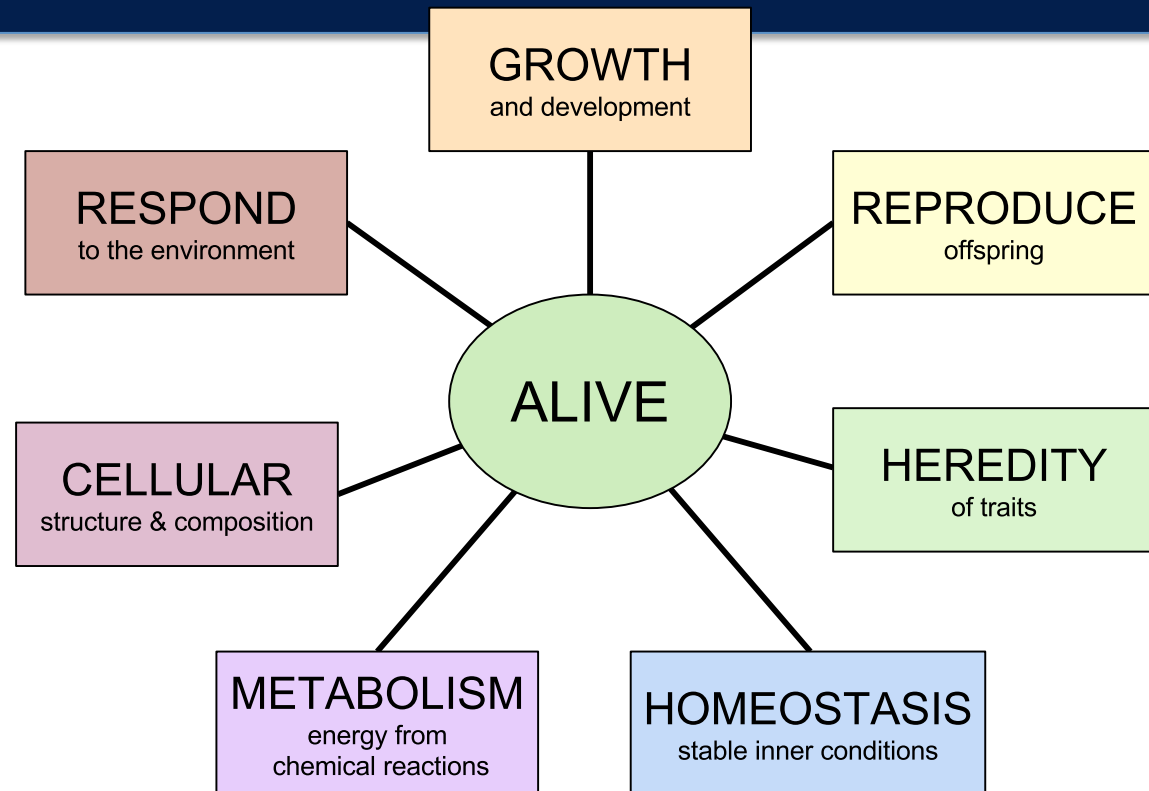
Are we being invaded by or coalescing with an alien life form?

The Human Side of This Coevolution
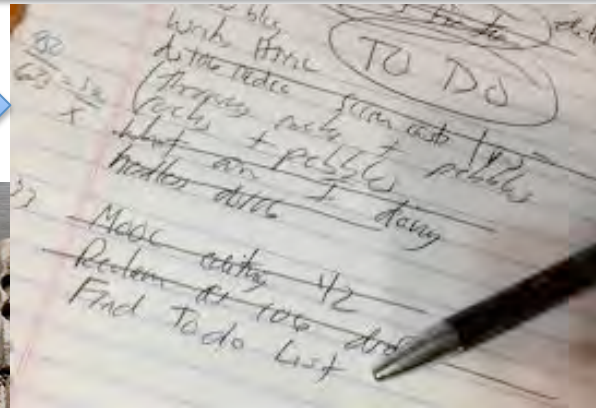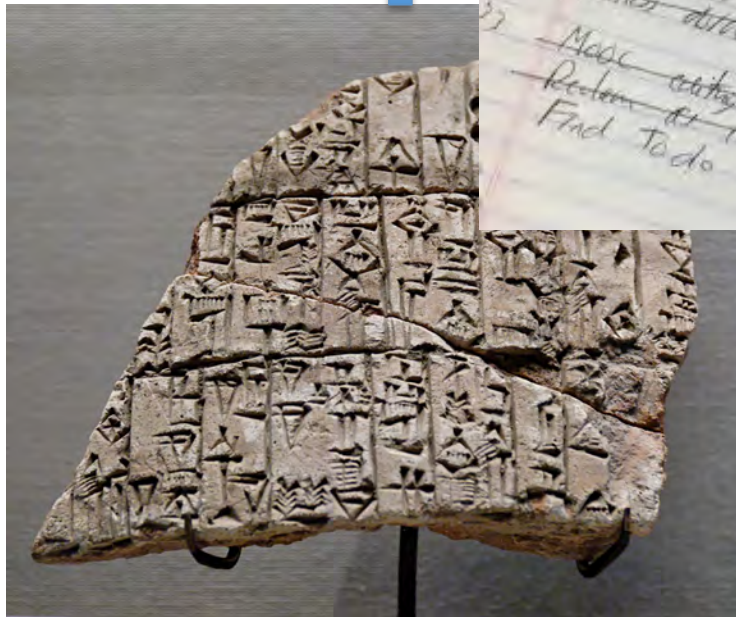
Intellectual Prosthetics
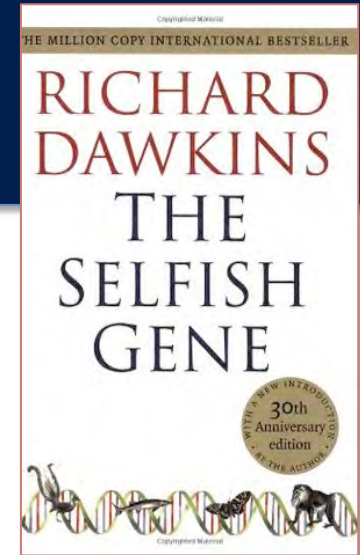
It is human culture and cognition ("memes" per Dawkins) that are coevolving, not (yet) biology.

Reproduction? Heredity? Mutation?

[Photo by Max Pixel, released to public domain - CC0]

Symbiosis

"If computers and software form organisms, then they depend on us for their procreation. We provide the husbandry and serve as midwives. In exchange, we depend on them to manage our systems of finance, commerce, and transportation. But more interestingly, the machines make the humans more effective at the very husbandry that spreads the software species.

….

the software survives and evolves only if the company survives and evolves, and vice versa."

The Creative Partnership of Humans and Technology

PLATO AND THE NERD

EDWARD ASHFORD LEE

19
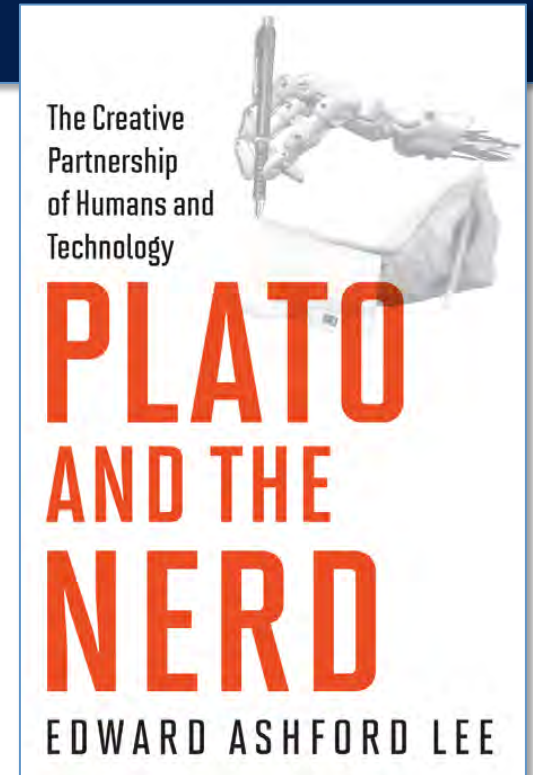
Will We Become Cyborgs?

We are already integrating technology into our biology.

CVCINA

By Unknown Master, Italian (1570s)
Web Gallery of Art, Public Domain

Quando a l'huome insatiabile dispiacque
Le giande usar per cibo è nutrimento.
Questa figura nua nel mondo nacque
A diletar il gusto a suo contento.

Onde non più la purità dell'acque
Cadé nel variato condimento.
Ma mille sughi d'herbe e mille odori
Chi all'appetito dan mille sapori

Obligate Endosymbiosis

Lynn Margulis (1938-2011)
[Photo by Jpedreira, CC BY-SA 2.5]
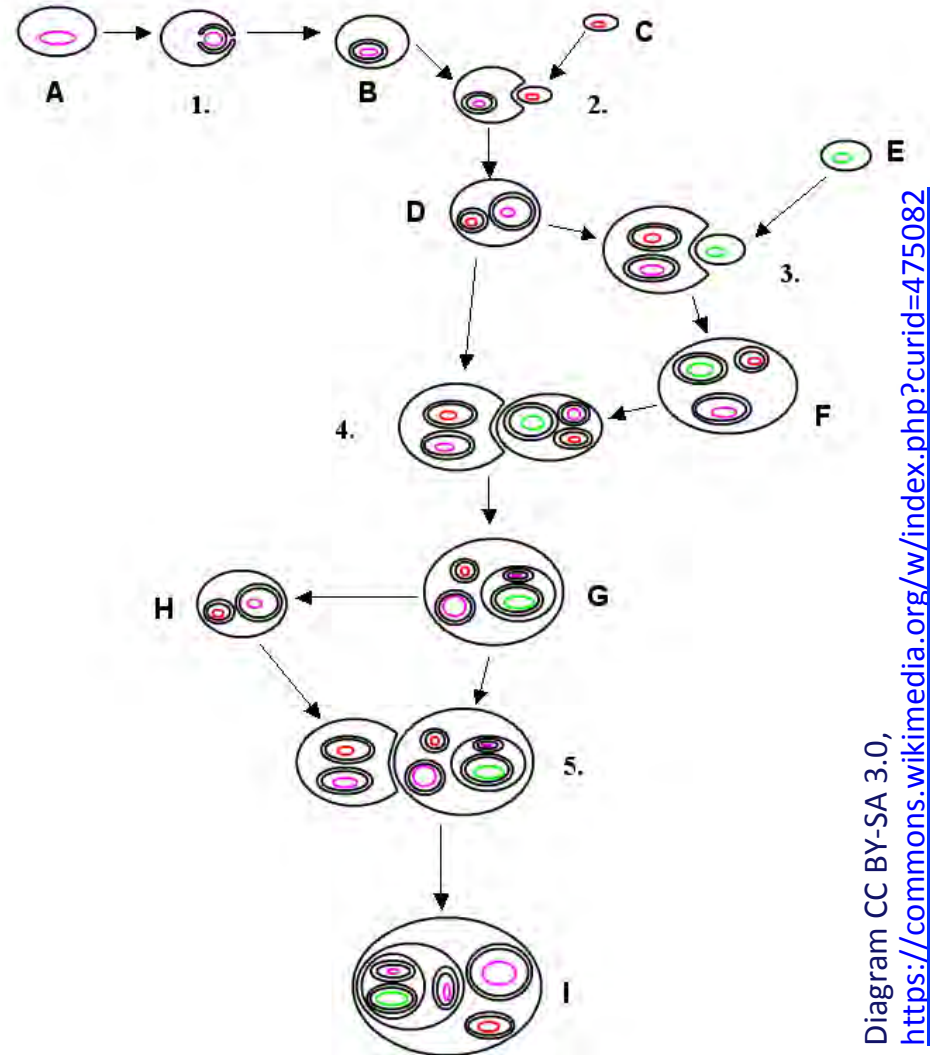
So, Are They Alive?

This depends on what you mean by "alive," but there is no doubt they share many features with biological beings.

And more importantly, their relationship with us is much like a biological symbiosis.

Three Questions

1. Are we going to lose control of them?
2. Are they alive?
3. Are they going to match and exceed us?

Computers already exceed us in many dimensions.
So the interesting question is: will they match us?

Are We Digital?

HAL, the computer in Stanley Kubrick's 1968 movie, *2001: A Space Odyssey*

Freeing the Mind From Matter

- Are we alone?

- Teleportation?

- The singularity?

- Uploading?

LIFE 3.0

BEING HUMAN IN THE AGE OF
ARTIFICIAL INTELLIGENCE

MAX TEGMARK
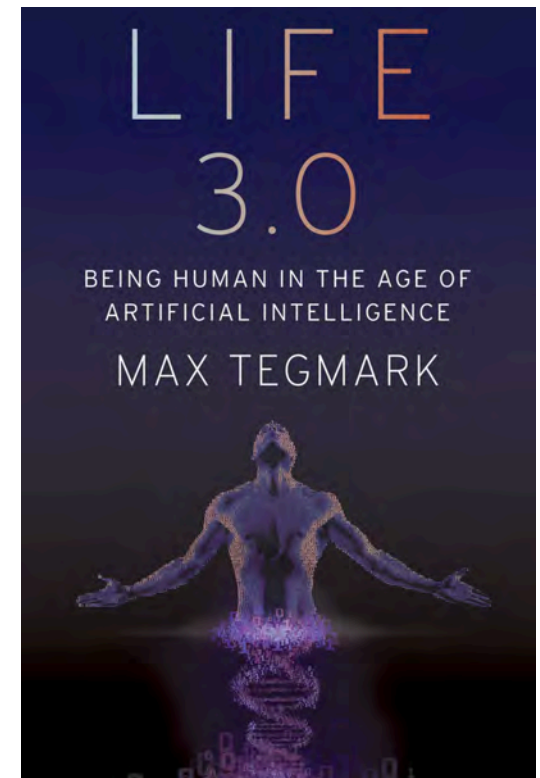
Teleportation and Uploading

What happens to "I"?

- Is the reconstruction the same "I"?
  - How can we tell?

- What if the original is not destroyed?
  - Two "I"s?

- What if a backup copy is later instantiated?
  - Two "I"s of different ages?

The Sense of Self Per Three Philosophers

What happens to "I"?

- Derek Parfit:
    - The notion of "I" makes no sense.

- Daniel Dennett:
    - "I" is a fiction, an illusion, a social construction.

- Douglas Hoftstadter
    - "I" can be in two places at once.

A Simpler Answer: "I" Is Not Digital
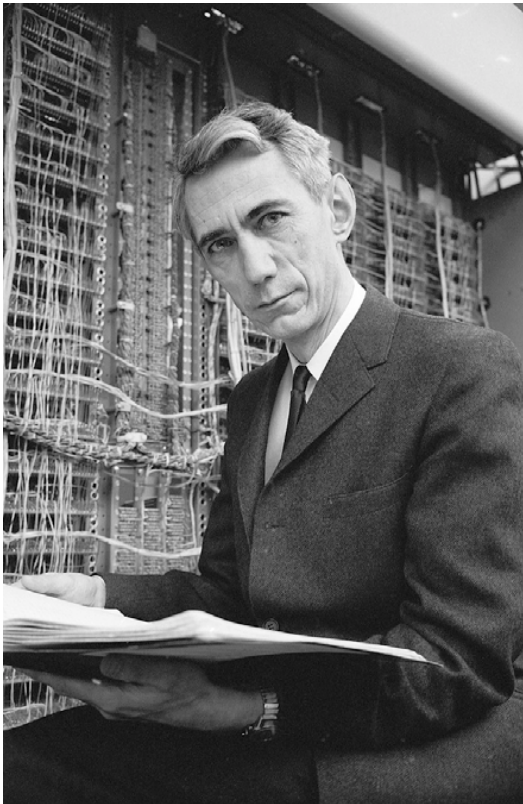
Claude Shannon

Shannon showed in 1948 a noisy channel can, in principle, perfectly convey a finite number of bits (the "channel capacity").

The converse is even more important: A noisy channel *cannot* convey more than a finite number of bits.

28

The "Universal Machine" Fallacy

Turing machines:

- Algorithmic
- Digital
- Terminating

Alan Turing
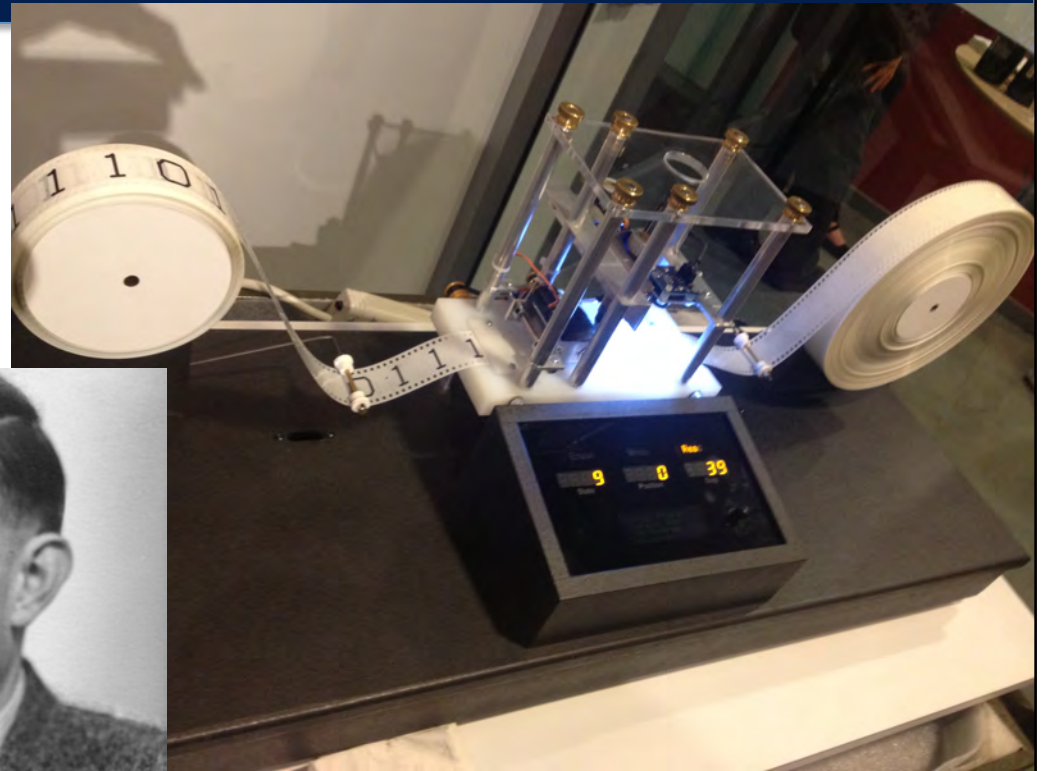
Machine designed by Mike Davey
[Photo by GabrielF, CC BY-SA 3.0]

A Universal Turing Machine is *Not* a Universal Machine

A machine that is (probably) not modeled in any useful way by a Turing machine.

It is neither digital nor algorithmic.

By Piotrus, CC BY-SA 3.0, via Wikimedia Commons

Cognitive Beings and Digital Abstractions

A cognitive being is a physical system.

A digital & computational system is an abstraction.
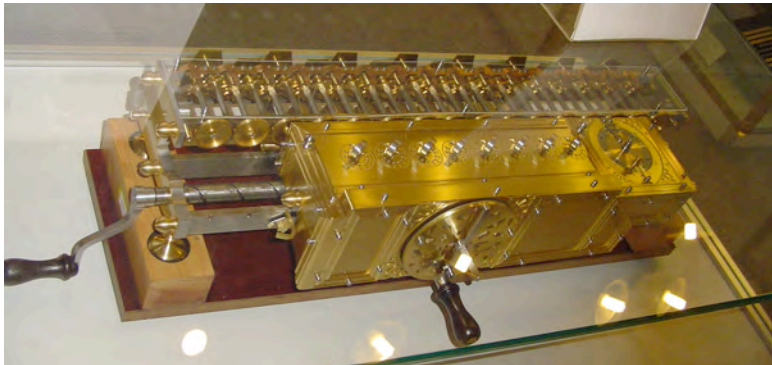
"Stepped reckoner" integer calculator designed by Leibnitz in 1672. [By Kolossos CC-BY-SA 3.0 via Wikimedia Commons]

GODEFROI GUILLAUME LEIBNITZ,
Né le 3 Juillet 1646 mort le 14 Novembre 1716.

31

How Many Machines?

My question:

Of all machines realizable in the physical world, how many are usefully modeled as digital and algorithmic machines?

Cantor and Cardinality

The smallest of all infinite sets are the countable sets. Elements of a countable set can be put into a one-to-one correspondence with the natural numbers:
0, 1, 2, 3, 4, …
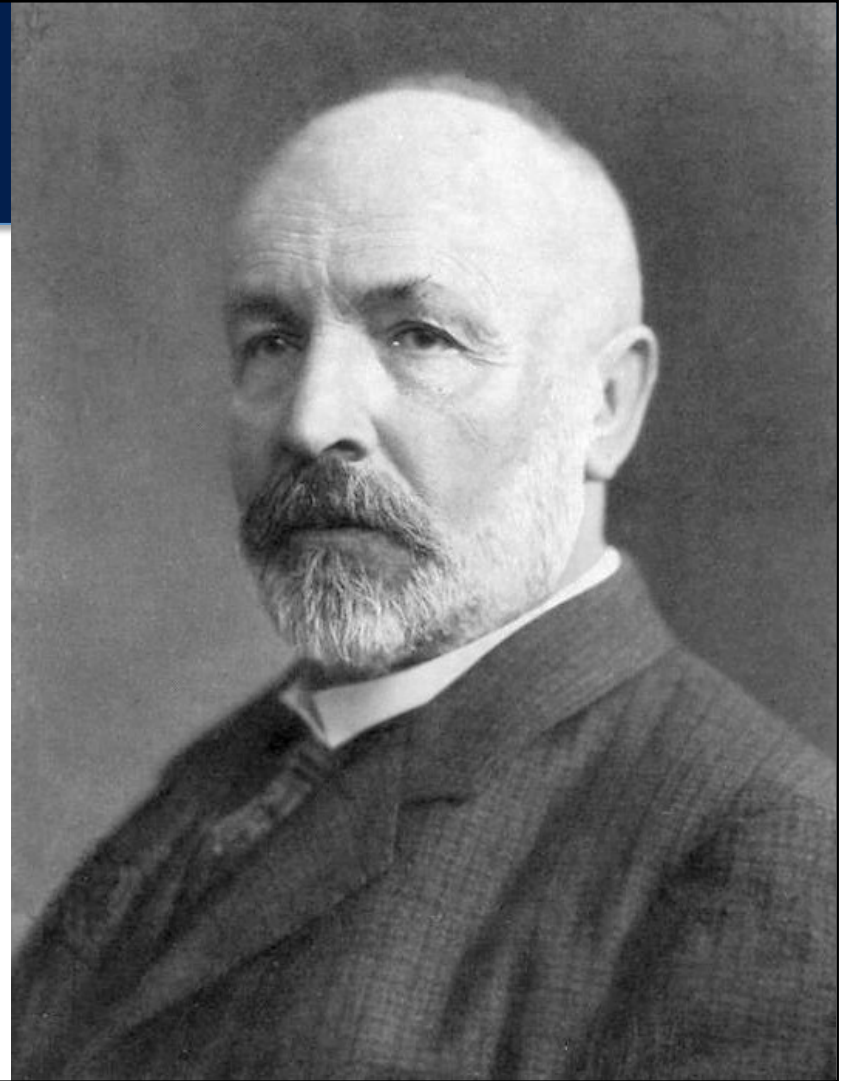
Georg Cantor (1845-1918)

The Set of Turing Machines is Countable

By Wvbailey CC BY-SA 3.0
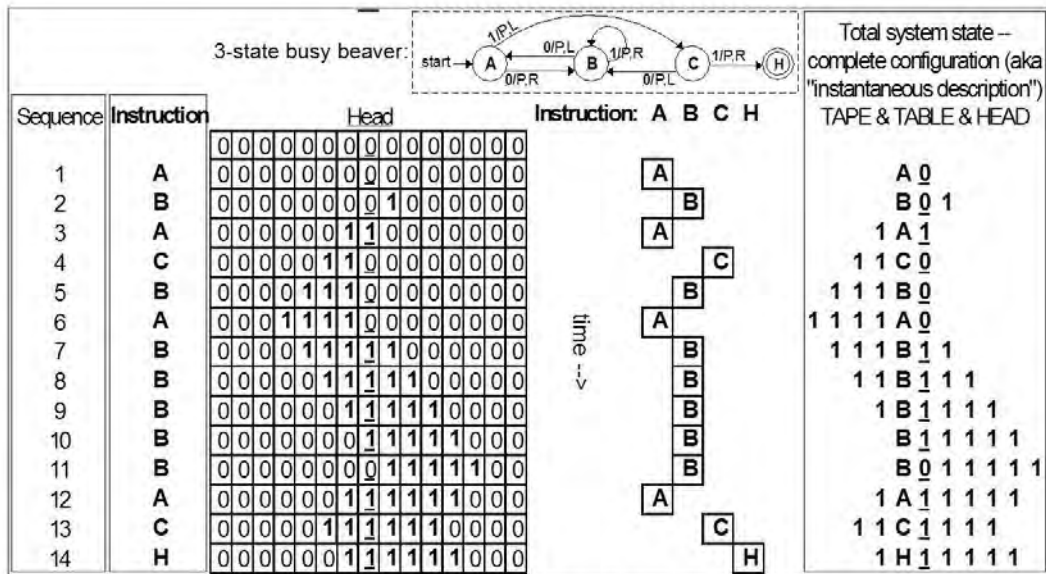
Progress of the computation (state-trajectory) of a 3-state busy beaver
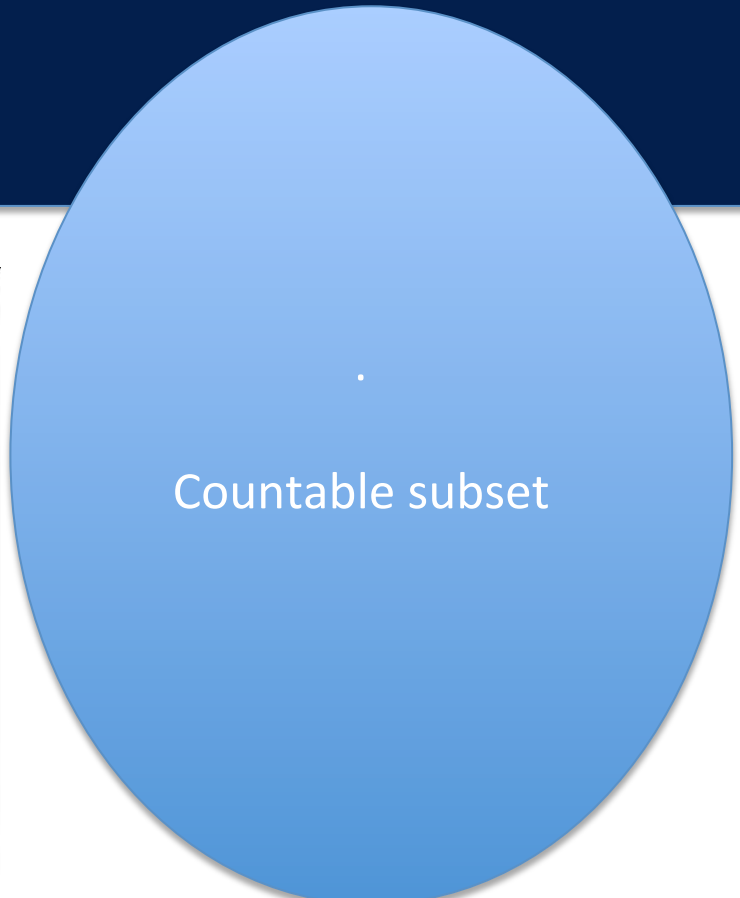
Countable subset

Uncountable set
E.g. set of all ODEs

34

The Set of Machines in Nature is (Probably) Not Countable

It would be truly remarkable if all machines in nature were Turing machines.

Not countable if:

- Time a continuum
- Space a continuum

Many physicists today say that physical world *is* countable.

Alfred A. Knopf,
2006
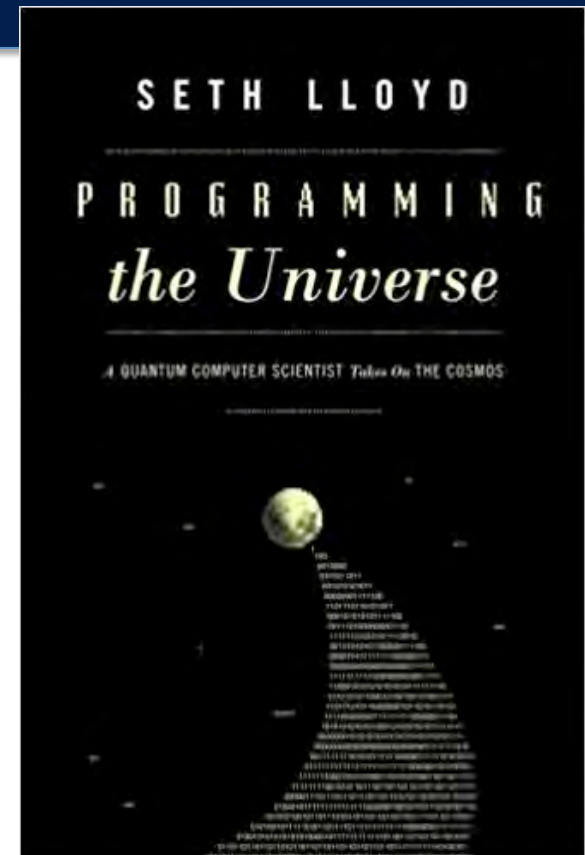
Or Maybe *Everything* is Digital?

Variants of the "Digital Physics" hypothesis:

1. The number of possible states of a physical system is finite.
2. Physical processes are digital and algorithmic.
3. Every physical process is a Turing computation.
4. The physical world is a computer.
5. The physical world is a simulation.

John Archibald Wheeler
"It from bit"
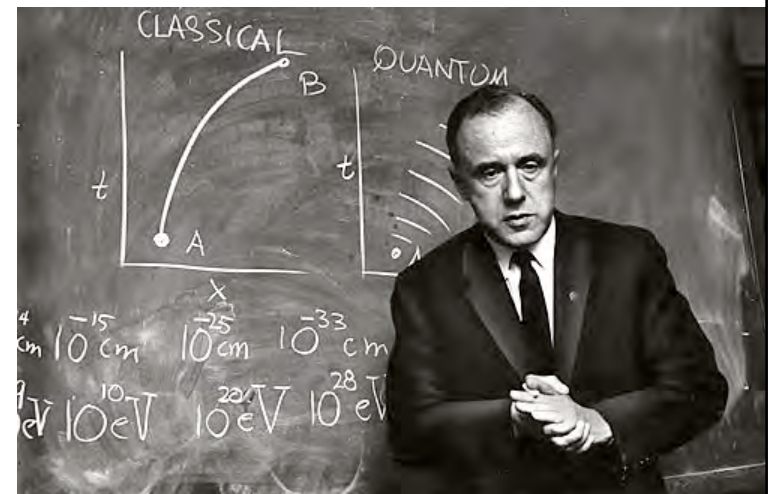
The Bekenstein Bound: One Justification for Digital Physics

The entropy in a finite region of space with finite energy is finite.

Stephen Hawking   Jacob Bekenstein

This result came from reconciling the second law of thermodynamics with the fact that black holes seem to swallow up entropy.

An Erroneous Deduction

Many physicists conclude that because entropy is bounded for a physical system with finite energy, that its state can be encoded completely with a finite number of bits.

This deduction is based on flawed mathematics!
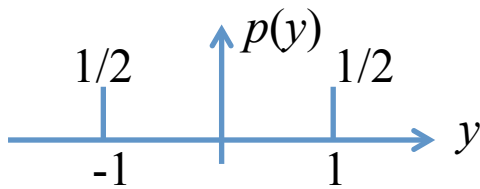
What is Information?
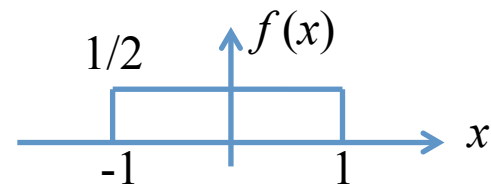
Information: Resolution of alternatives.

Discrete random variable

An observation can be encoded with one bit.

$$E(Y) = -\sum_{y \in \Gamma} p(y) \log_2(p(y))$$

Continuous random variable

No finite number of bits suffices to encode an observation.

$$H(X) = -\int_{\Omega} f(x) \log_2(f(x)) dx$$

Entropy is the expected information in an observation.

Discrete Entropy vs. Continuous Entropy

It is an error to conclude that finite entropy implies that the state can be encoded in a finite number of bits.

$$E(Y) = -\sum_{y \in \Gamma} p(y) \log_2(p(y))$$

Discrete entropy

$$H(X) = -\int_{\Omega} f(x) \log_2(f(x)) dx$$

Continuous entropy

For a gentle introduction to entropy and information, see Lee, *Plato and the Nerd*, MIT Press, 2017, chapter 7.

# Digital Physics is not a Scientific Hypothesis

The hypothesis that the state of a physical system can be encoded with a finite number of bits is not falsifiable by experiment.

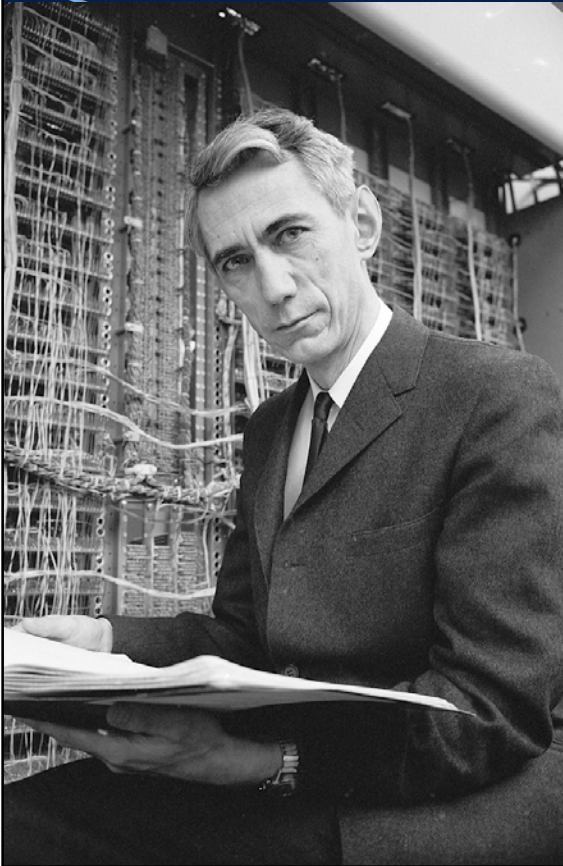Therefore, this hypothesis is not scientific (per the philosophy of Karl Popper).

It can only be taken on faith.

The Creative Partnership of Humans and Technology

PLATO AND THE NERD

EDWARD ASHFORD LEE
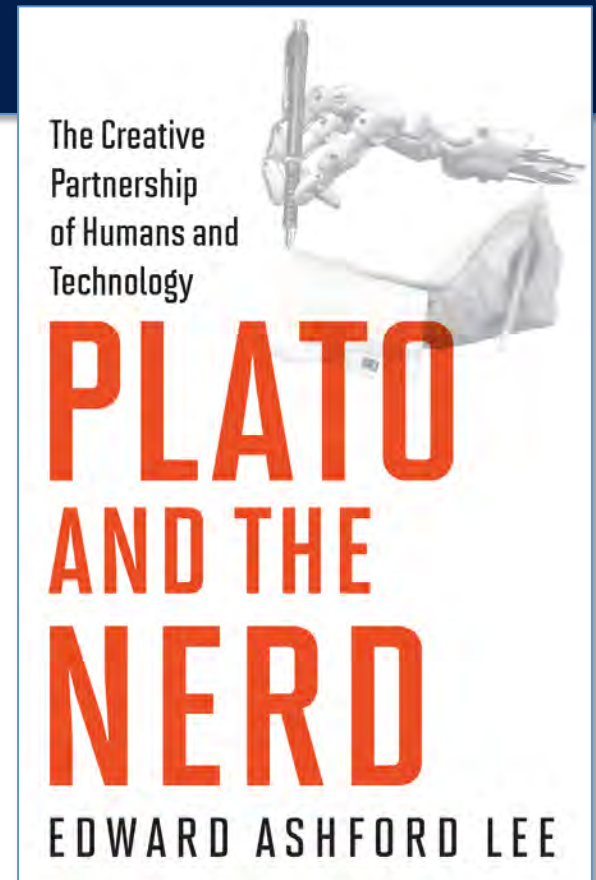
MIT Press, 2017

The Shannon Channel Capacity Theorem (Again!)

Shannon showed in 1948 a noisy channel can, in principle, perfectly convey a finite number of bits (the "channel capacity"), but no more.

Every measurement of the physical world is noisy (unless you first assume that digital physics is true).

Claude Shannon

What Are the Odds?

Is cognition digital
and computational?

We need
evidence…

Digital/
computational
processes

Processes in nature

Dataism is a Faith

Yuval Noah Harari

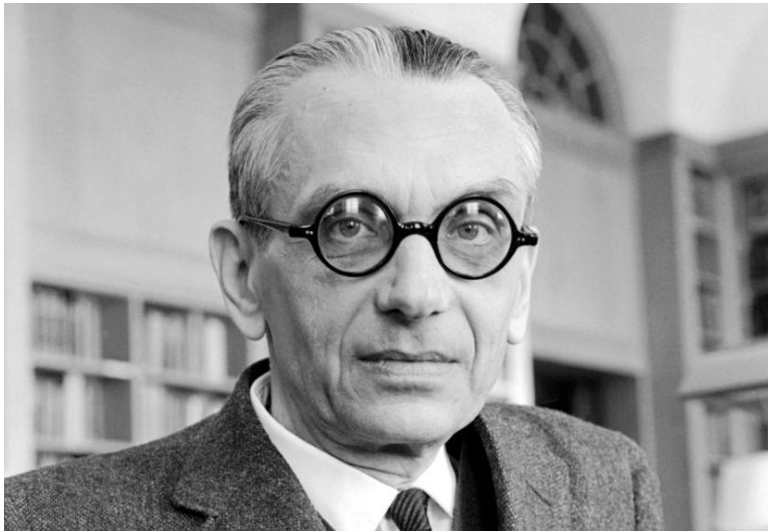[Photo By Daniel Naber
–CC BY-SA 4.0]

Possible Mechanisms in the Brain that are Beyond Digital/Computational

- Timing

- Interaction

- Chaos (induced by feedback loops)

- Nondeterminism

- Chemistry

- Embodiment

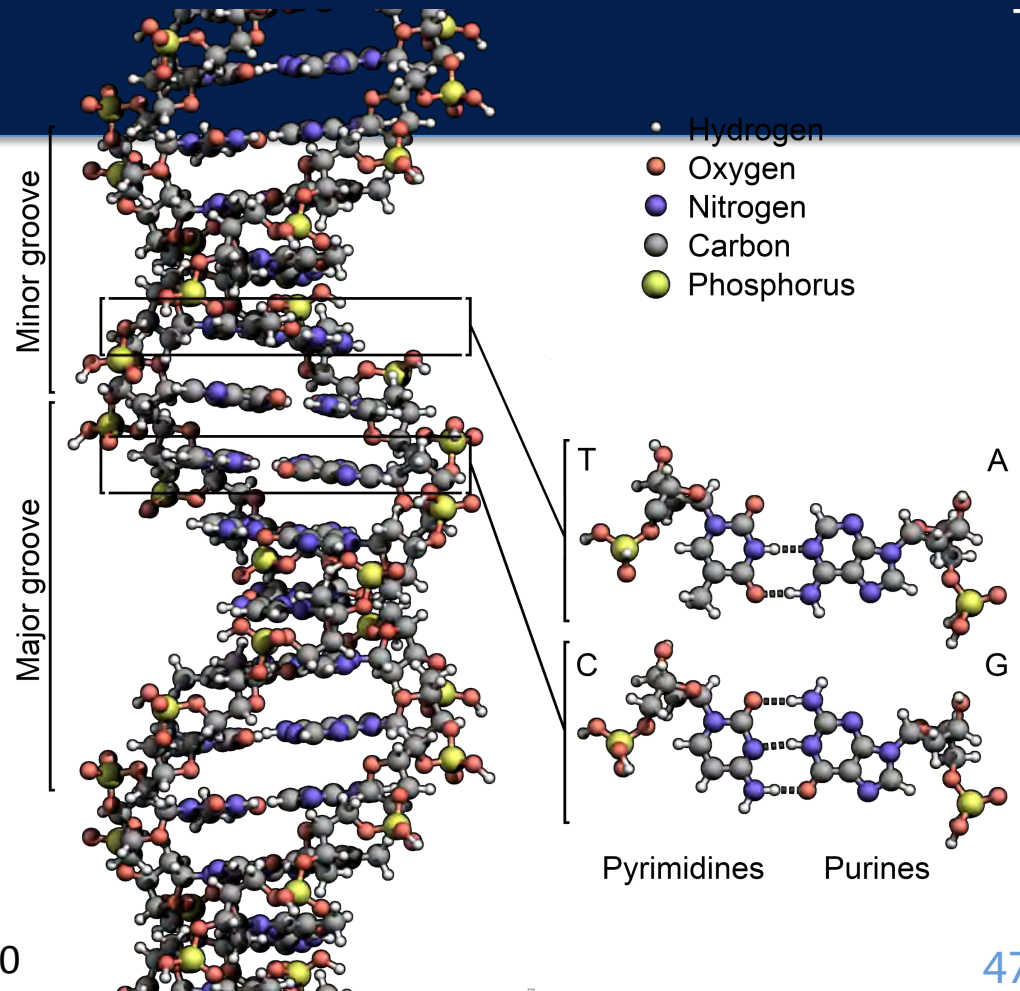# No Universal Machine Has Yet Been Invented

Kurt Gödel

If one is ever invented, it will not be, at its essence, a discrete, algorithmic, terminating process.

The DNA Fallacy

Every human alive today is the endpoint of continuous, unbroken, biological process dating back about four billion years.
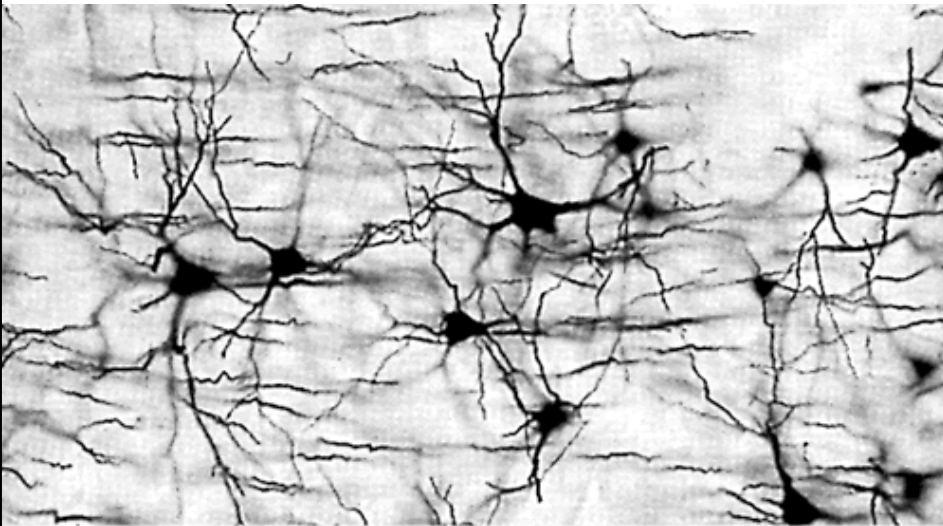

Minor groove

Major groove

Hydrogen
Oxygen
Nitrogen
Carbon
Phosphorus

T   A

C   G

Pyrimidines   Purines

The Connectomics Fallacy

- Neurons fire discretely. (McCulloch and Pitts, 1940s)
- Neurons combine to realize logic functions.
- Logic functions can be realized on other hardware (Putnam, 1960s).
- Connections will reveal brain function (Lichtman, 2000s).

Camillo Golgi's method (1870s) gives a misleading picture of the brain.

Connectomics:
A More Complete Picture of the Brain

Can we understand brain function by studying the wiring diagram, even in principle?

Jeff Lichtman, Harvard

If Cognition is not a Digital, Algorithmic Process, then…
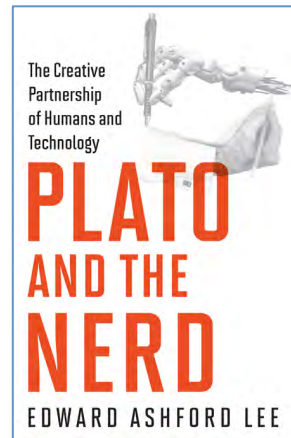
"Your mind is entirely your own."

Three Questions

1. Are we going to lose control of them? *No.*

2. Are they alive? *Maybe.*

3. Are they going to match and exceed us?
   They already exceed us, but they will never match us.

The Creative Partnership of Humans and Technology

PLATO AND THE NERD

EDWARD ASHFORD LEE

MIT Press, 2017

Living Digital Beings

Edward Ashford Lee

MIT Press, 2020