Building Artificial General Intelligence

Peter Morgan

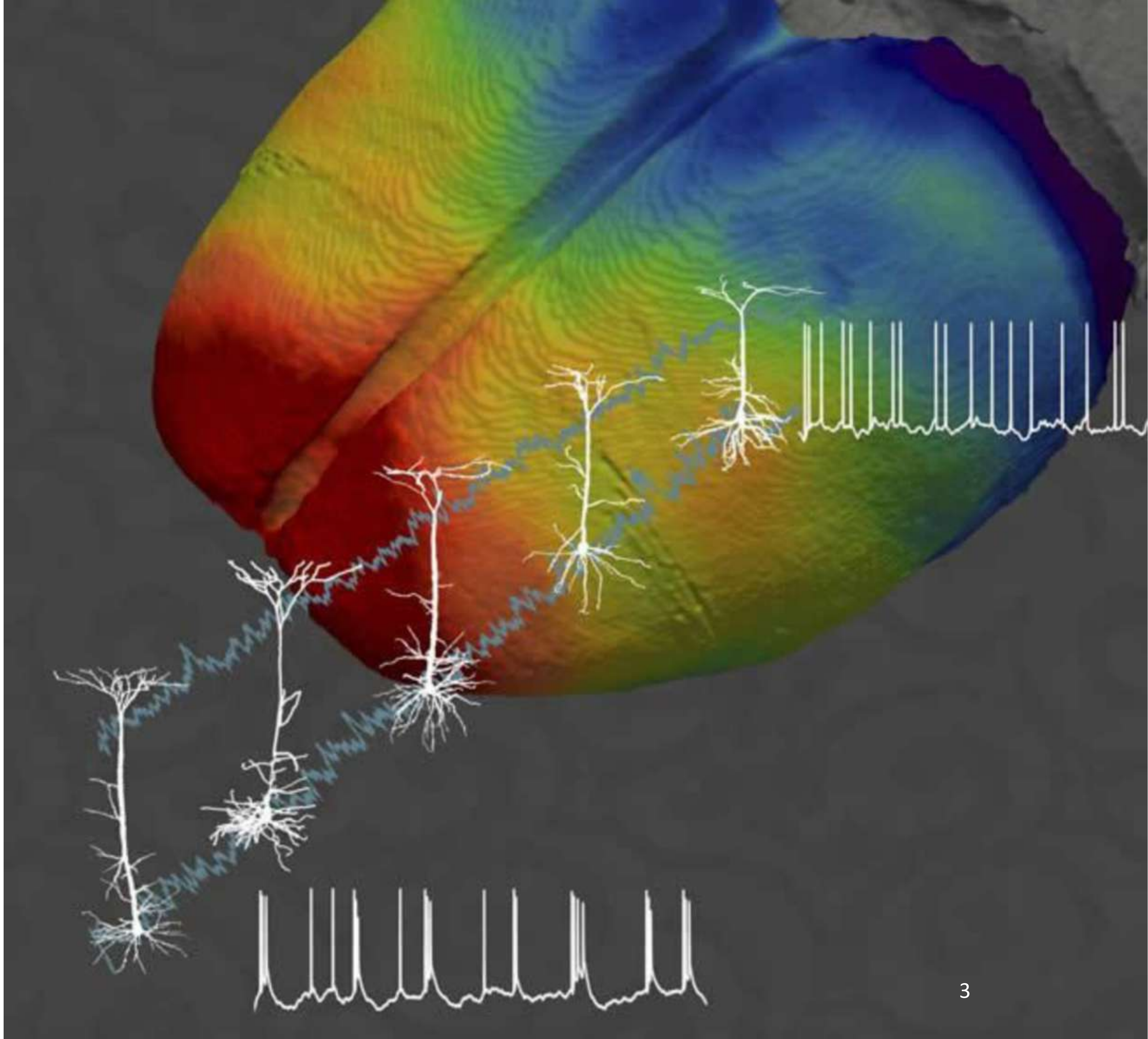
CEO TURING.AI

Outline of Talk

- What is Intelligence?
- Physical Systems
  - Biological
  - Non-biological
- Deep Learning
  - Recap
- AGI
  - Overview
  - Active Inference
  - Building AGI
- Conclusions

TL;DR

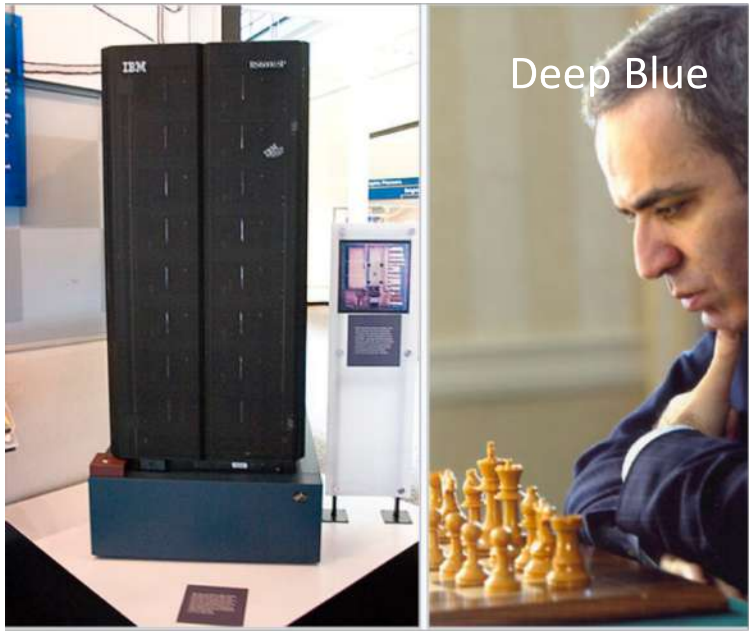
Using the physical principles of active inference, I believe we can build AGI systems over the coming years.



Motivation

Why do we want to build AGI?

- To solve (general) intelligence
- Then use it to solve everything else
  - Medicine, cancer, brain disease (Alzheimer's ...), longevity, physics, maths, materials science, social
- Naturally extend to superintelligent systems

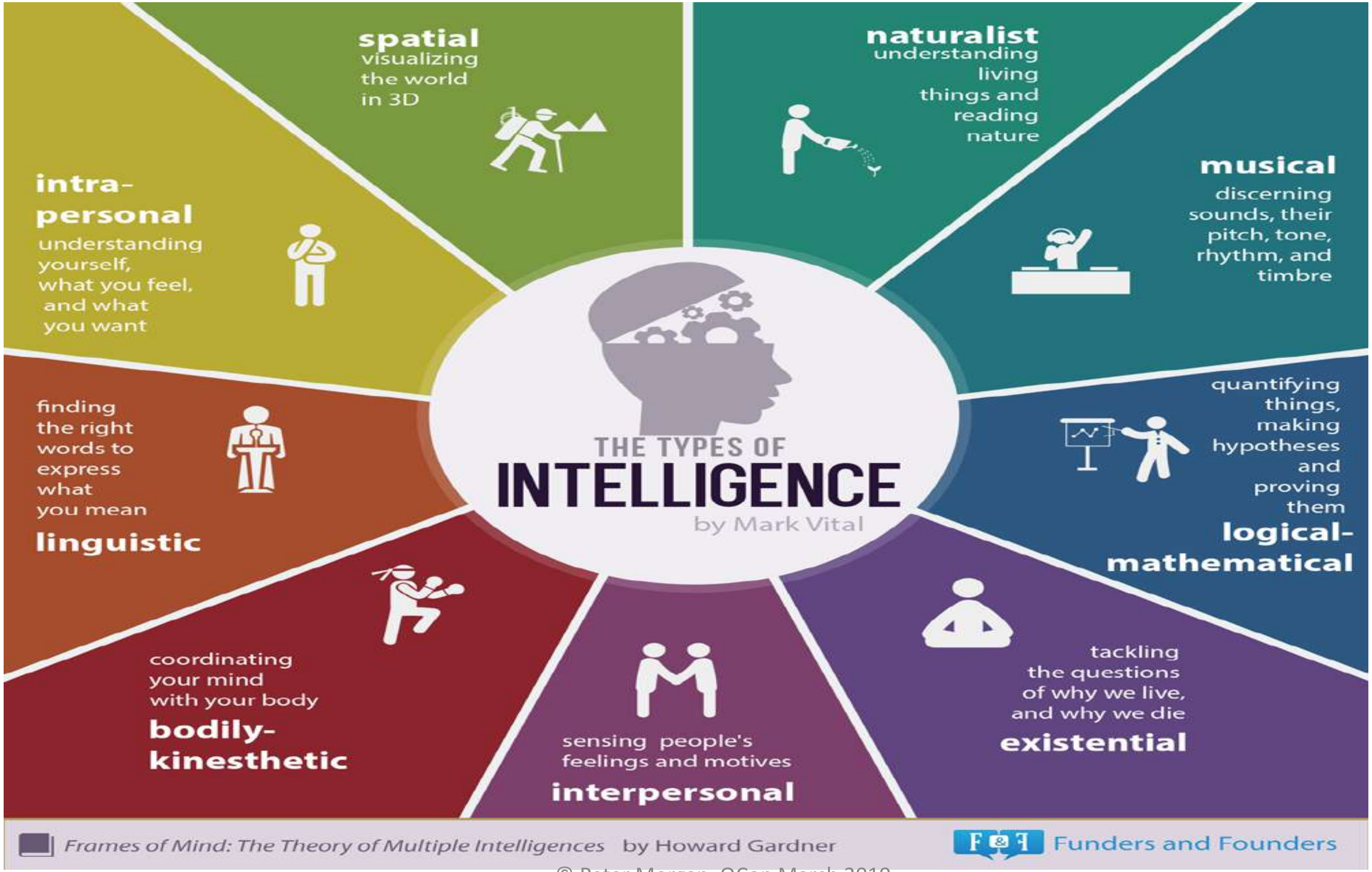


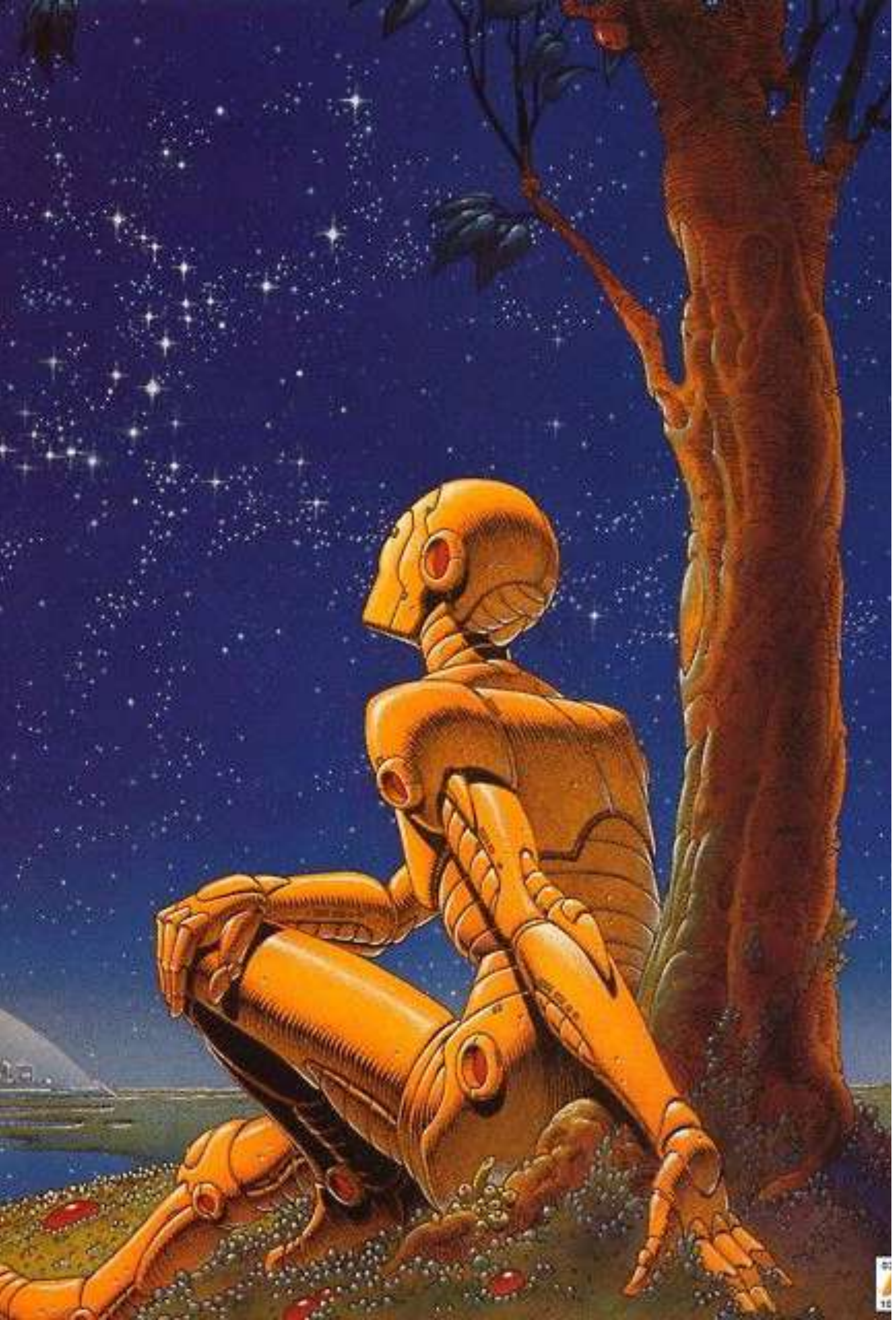
What is general intelligence?
(Ask the audience)

What is intelligence?

No matter how impressive,
these are all examples of
Narrow AI

What is Intelligence?



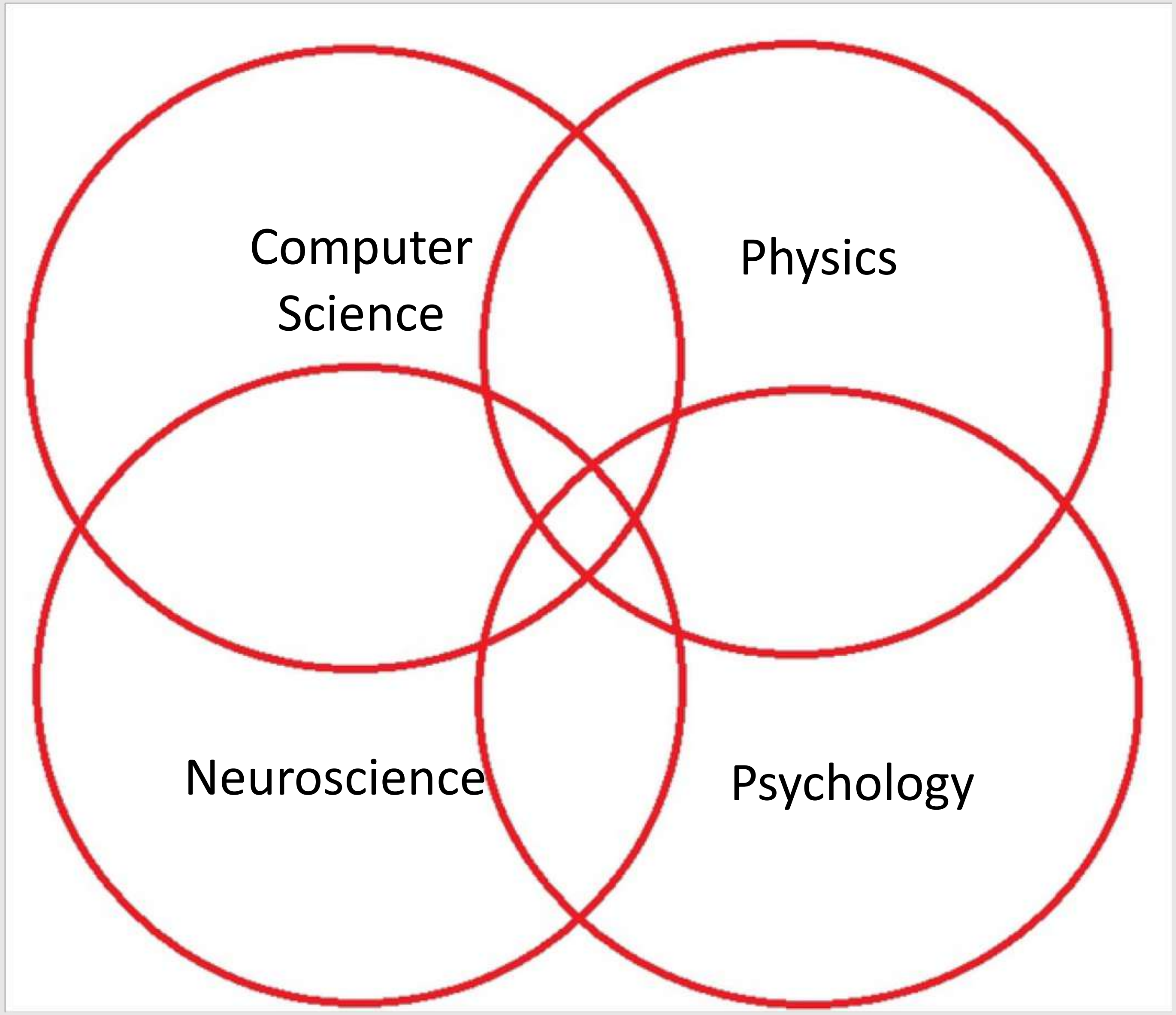


How far have we come?

- Logical-mathematical 50% (calculation not creativity)
- Linguistic 50% (statistical only)
- Spatial 50% (SLAM)
- Bodily-kinesthetic 30% (Atlas)
- Naturalistic 10%
- Musical 50% (Google Magenta project)
- Interpersonal 10%
- Intrapersonal 5%?
- Existential 0%

How will we get to AGI?

It takes a village to create an AGI

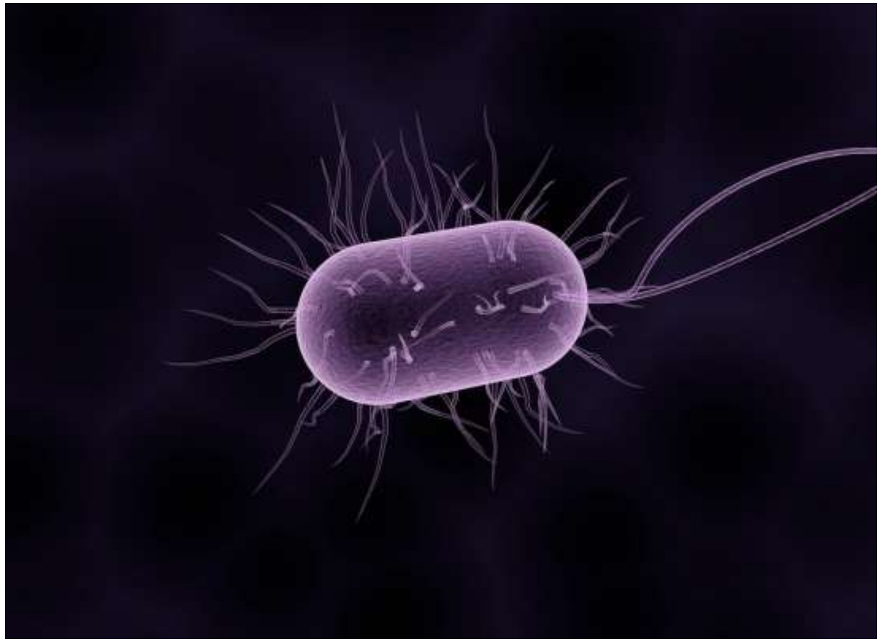


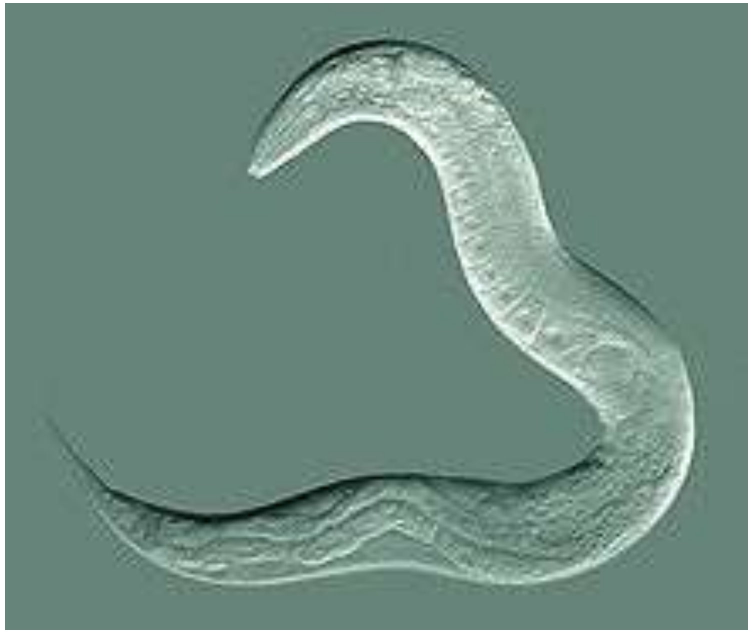
Outline

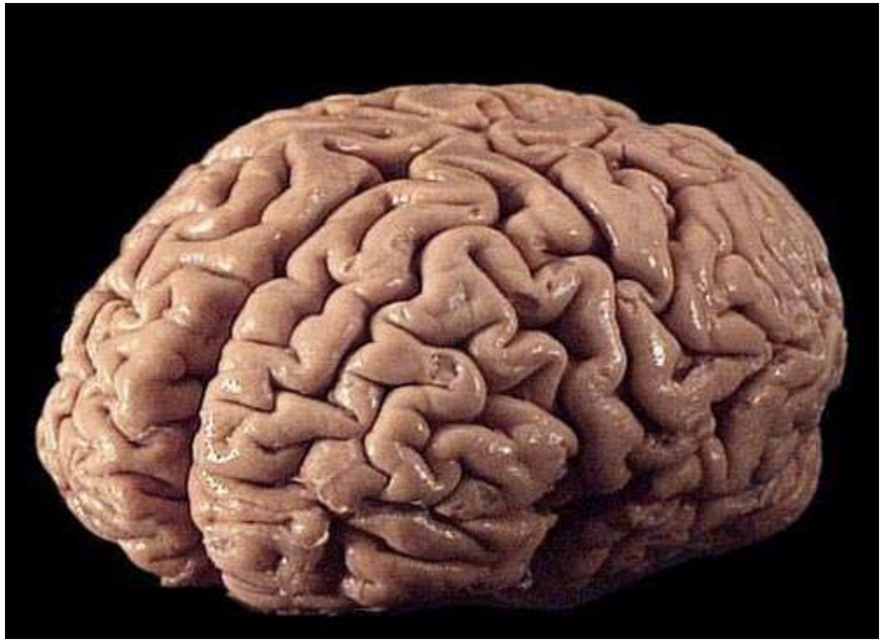
- What is Intelligence?
- **Physical Systems**
  - Biological
  - Non-biological
- Deep Learning
  - Recap
- AGI
  - Overview
  - Active Inference
  - Building AGI
- Conclusions

Intelligence in Physical Systems

- **Biological (neuroscience)**
  - Plants, bacteria, insects, mammalian
- **Non-biological (computer science)**
  - CPU, GPU, FPGA, ASIC
  - Neuromorphic
- **Quantum (physics)**
  - Quantum processors, algorithms
  - Biology?





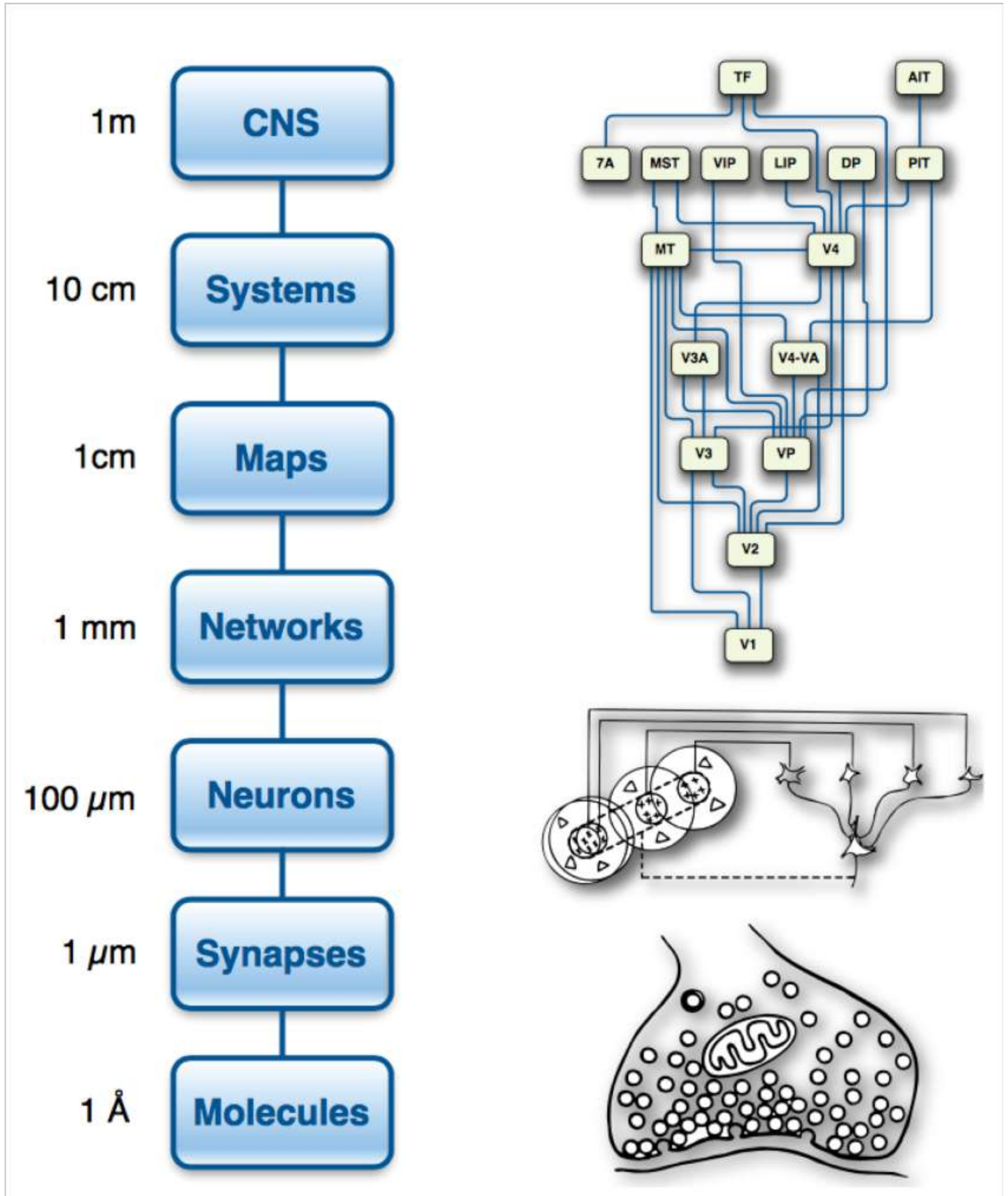


Biology

We would like to build human level intelligence

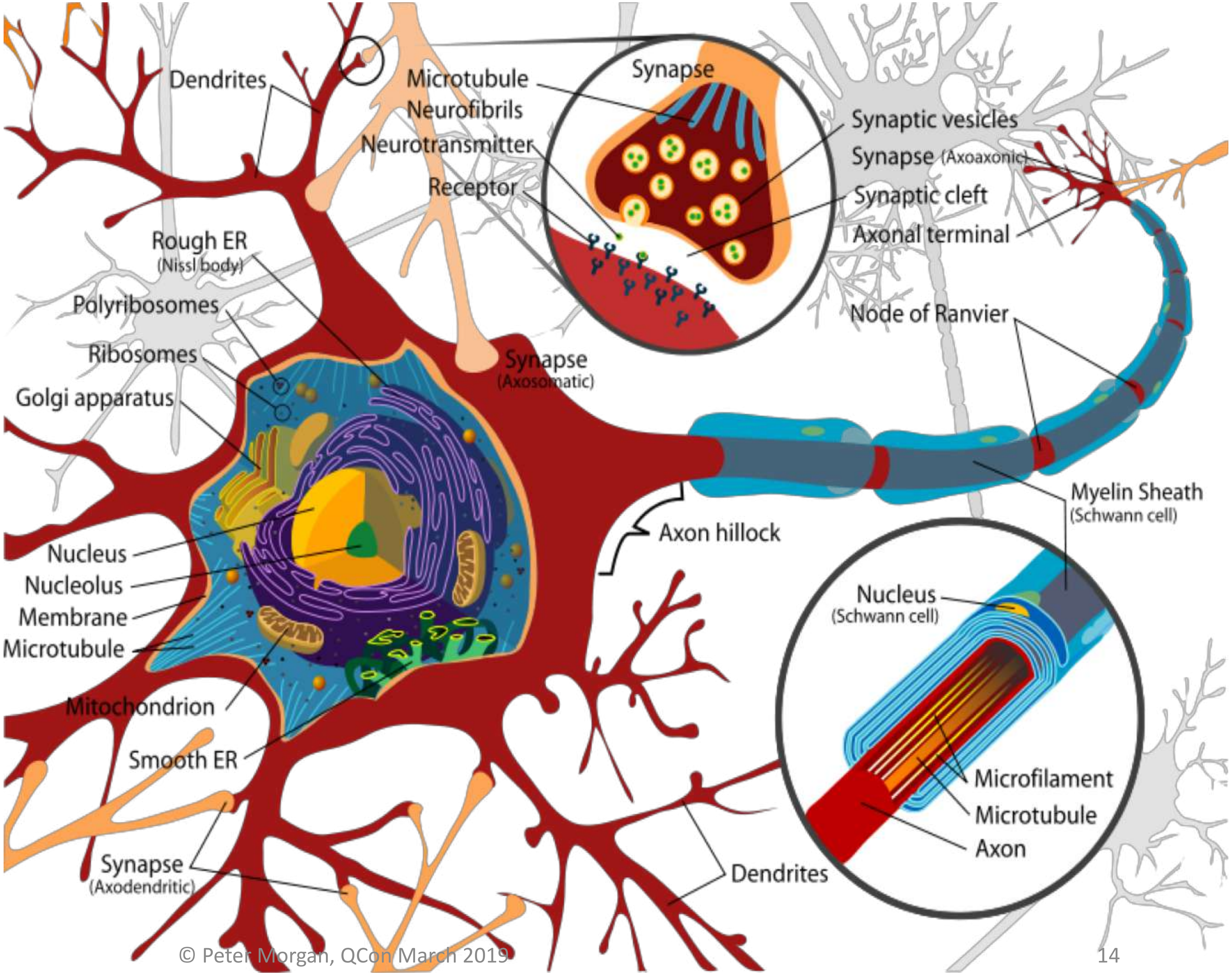
Biological Systems are Hierarchical

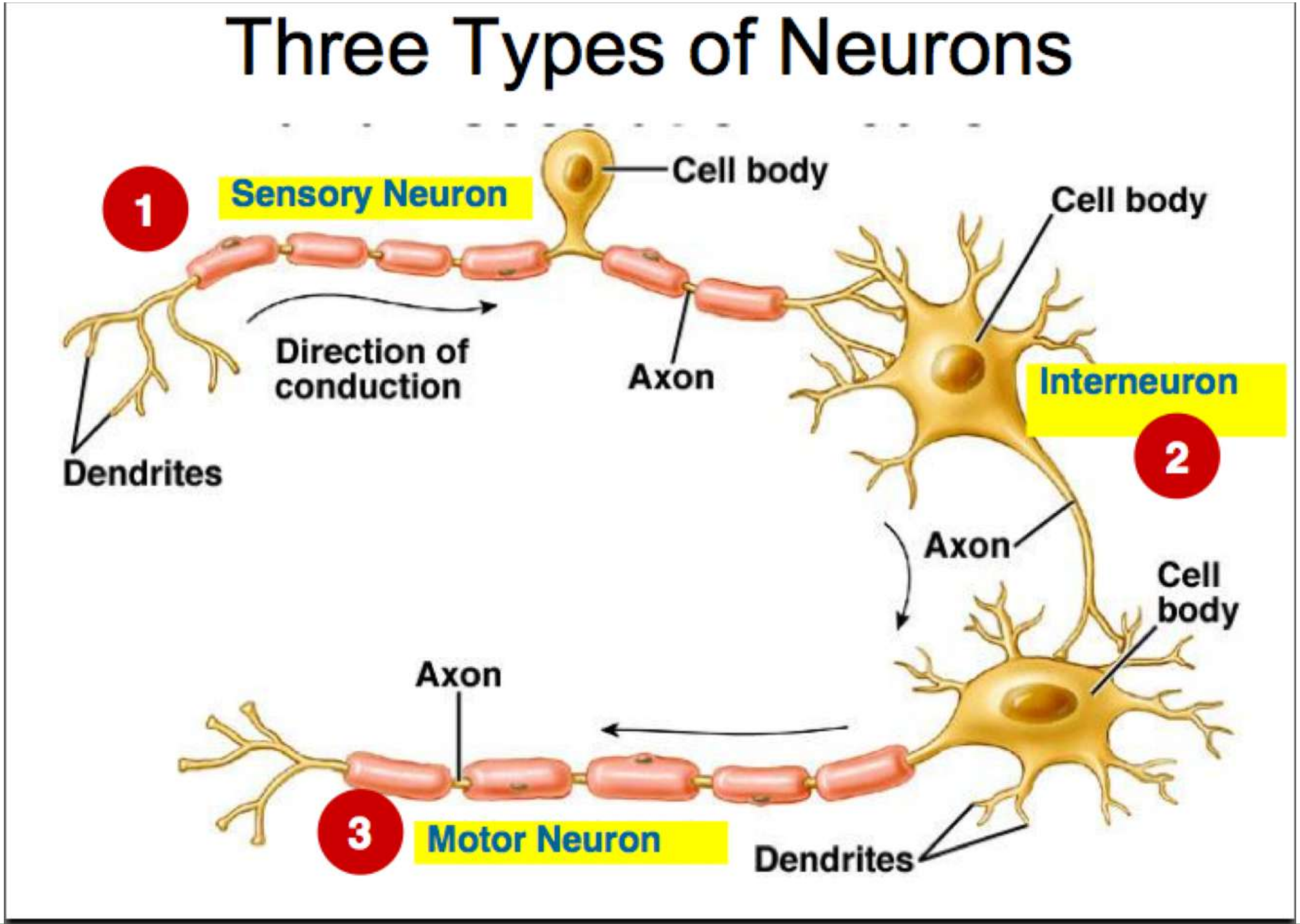
Intelligence is "emergent"



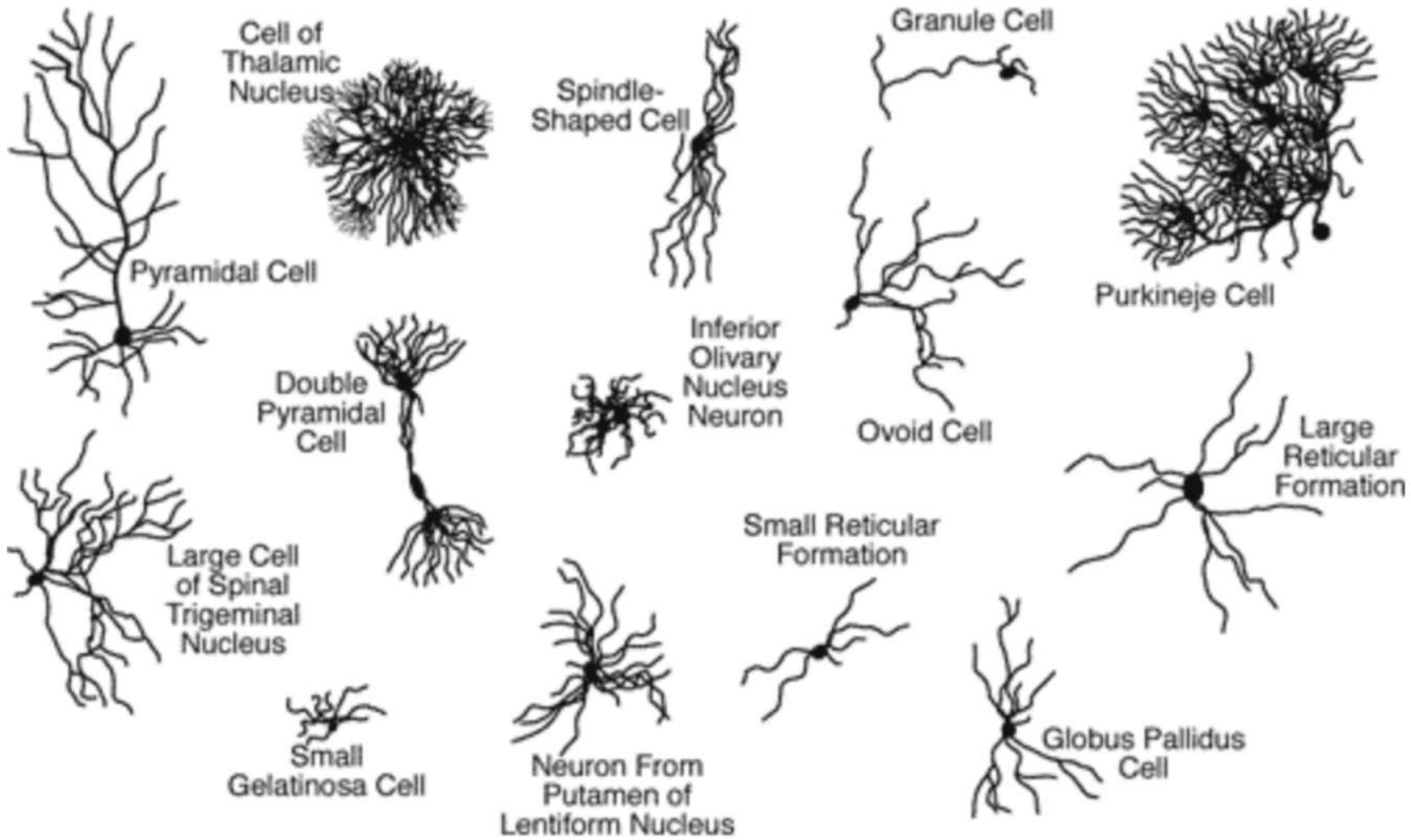
Biological Neuron Microstructure

Very complex




Three Types of Neurons
1
Sensory Neuron
Cell body
Cell body
Direction of conduction
Axon
Dendrites
Interneuron
2
Axon
Cell body
Axon
3
Motor Neuron
Dendrites

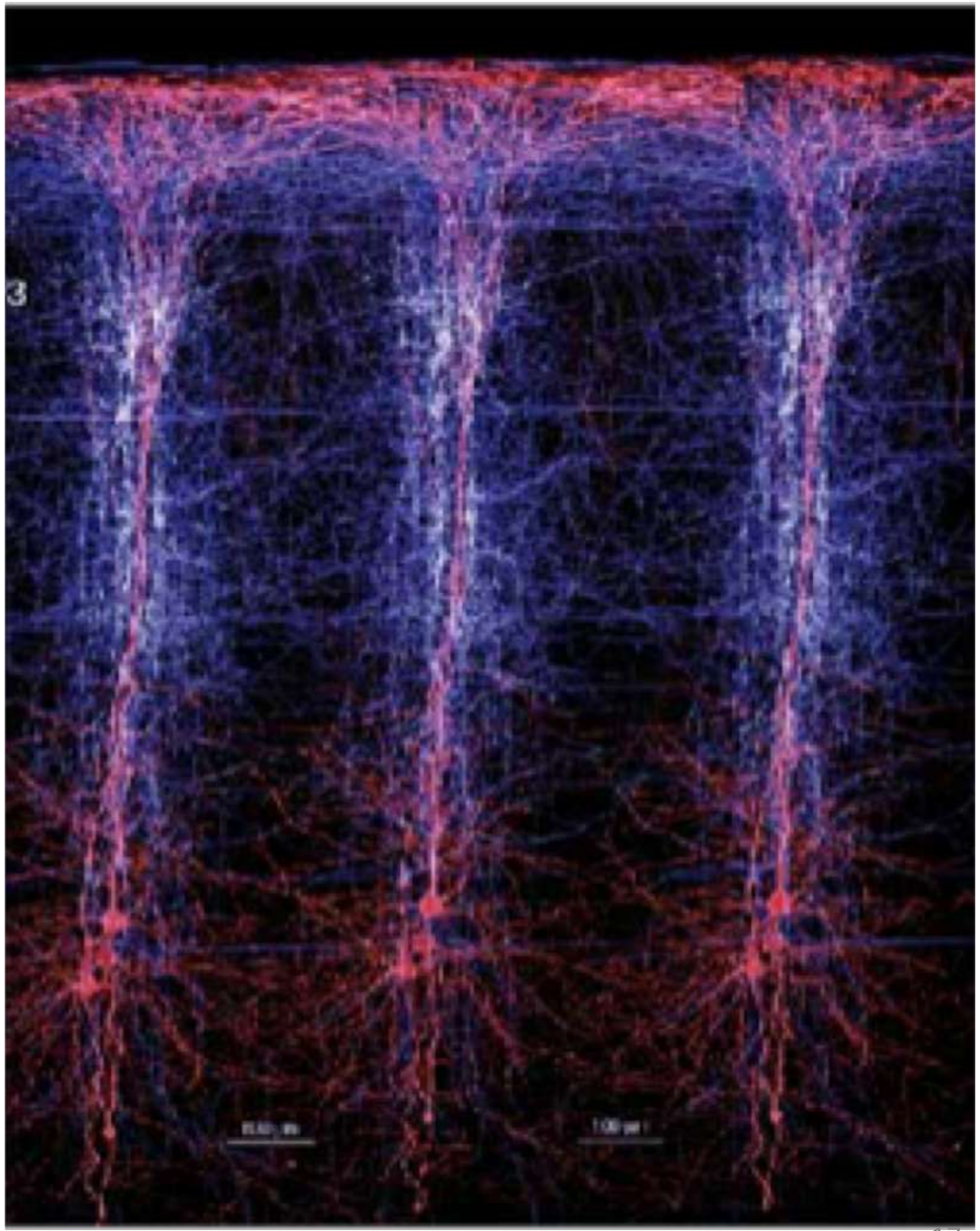
Different Morphologies



Cortical columns in the cortex

Approx. 2 million in human brain

Structure is repeated

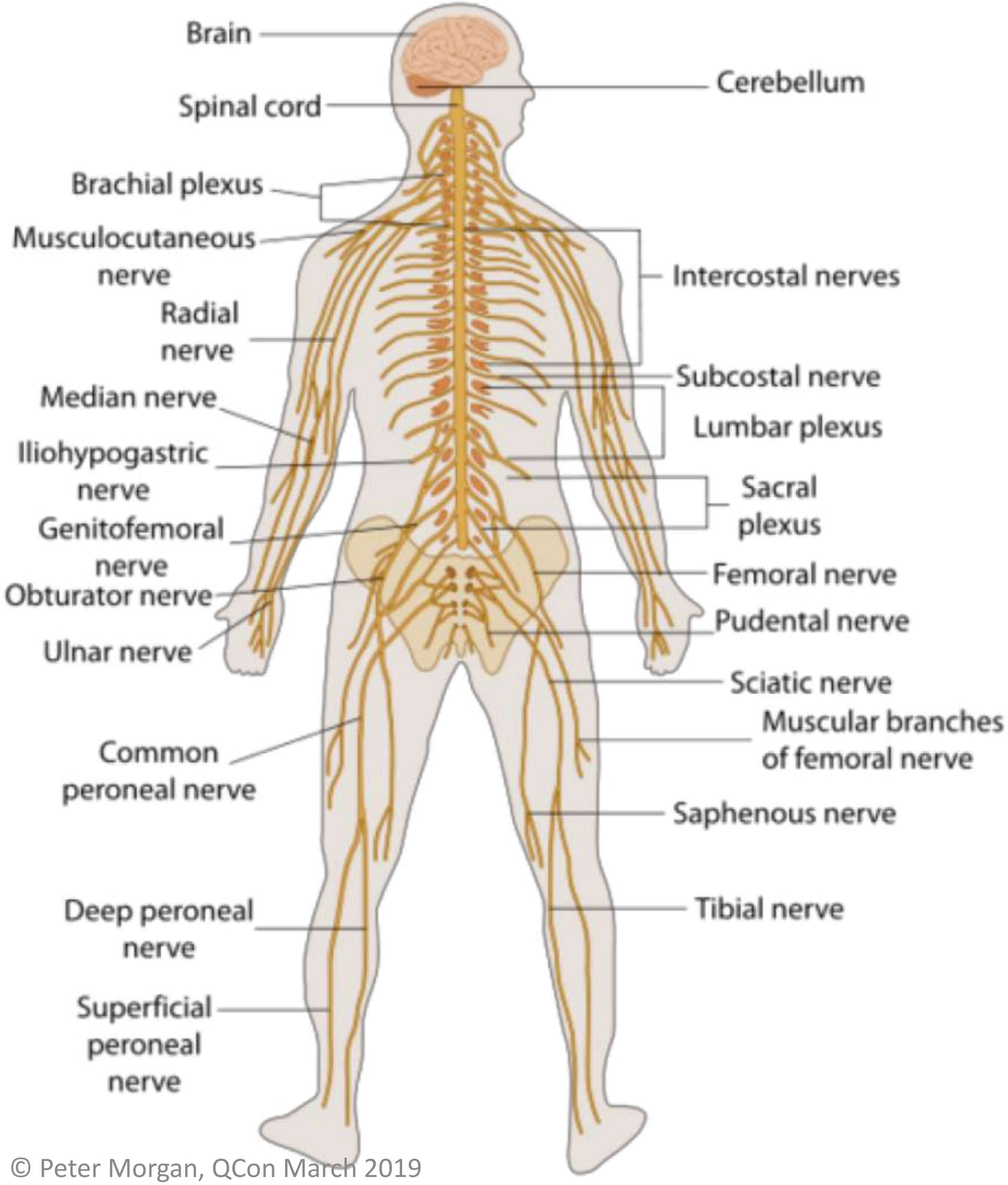


Human
Connectome

General
intelligence occurs
at this level

Central Nervous System (CNS)

Bodily intelligence



Social Systems

Cloud robotics
Swarm robotics

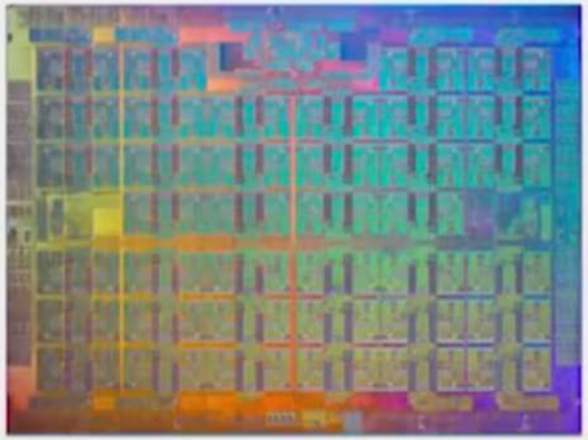
Non-biological Hardware

- **Digital**
  - CPU, GPU, FPGA, ASIC
- **Neuromorphic**
  - Various architectures
- **Quantum**
  - Different types

Digital Computing

- Abacus (mechanical, 2700 BCE)
- Charles Babbage (1830)
- Ada Lovelace
- Vacuum tubes (electronic, 1900)
- Alan Turing (1930's)
- Von Neumann
- ENIAC (1946)
- Transistor (Bardeen, Brattain, Shockley, 1947)
- Intel (1968)
- ARM (1990)
- Nvidia (1993)
- ASICs – TPU (2016)

Processor are built for specific workloads



CPU

Scalar

Designed for office apps

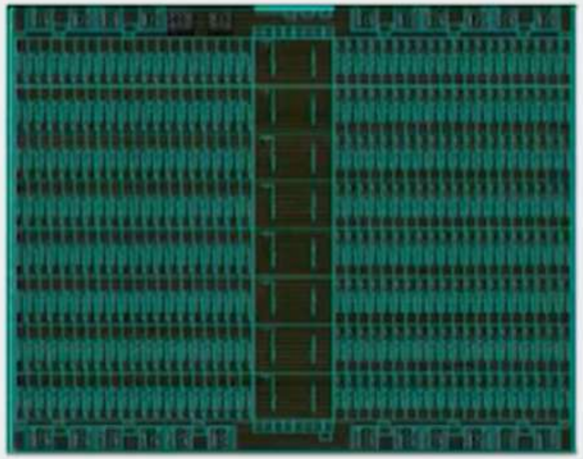
Evolved for web servers

GPU

Vector

Designed for graphics

Evolved for linear algebra



IPU

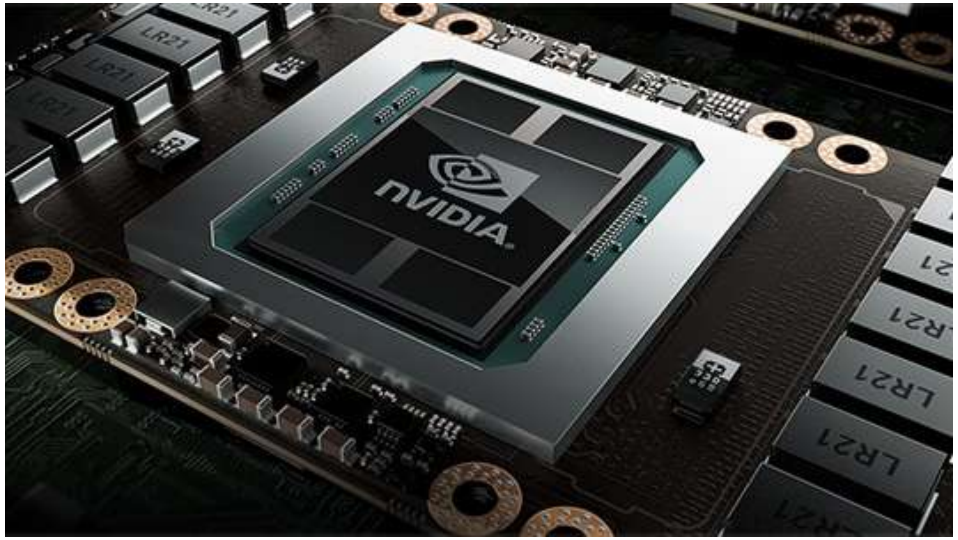
Graph

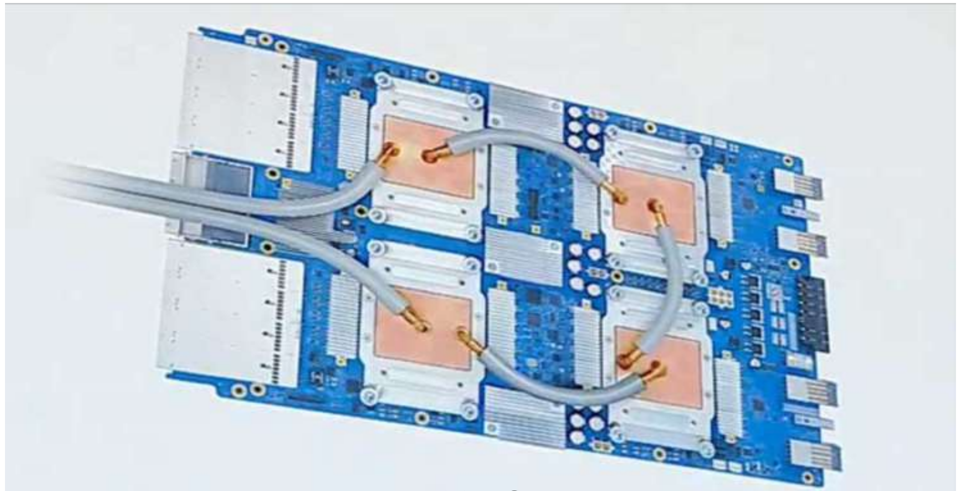
Designed for intelligence

Cray-1
1976

160 MFlops

Today





TPU v3

Graphcore IPU

Cloud TPU's

Over 100 PetaFlops - 1 billion Cray-1's!

Summit US
3 ExaFlops mixed precision
2 tennis courts area
250 Petabytes storage
13MW power
$200million
Announced 5 June 2018

No general intelligence

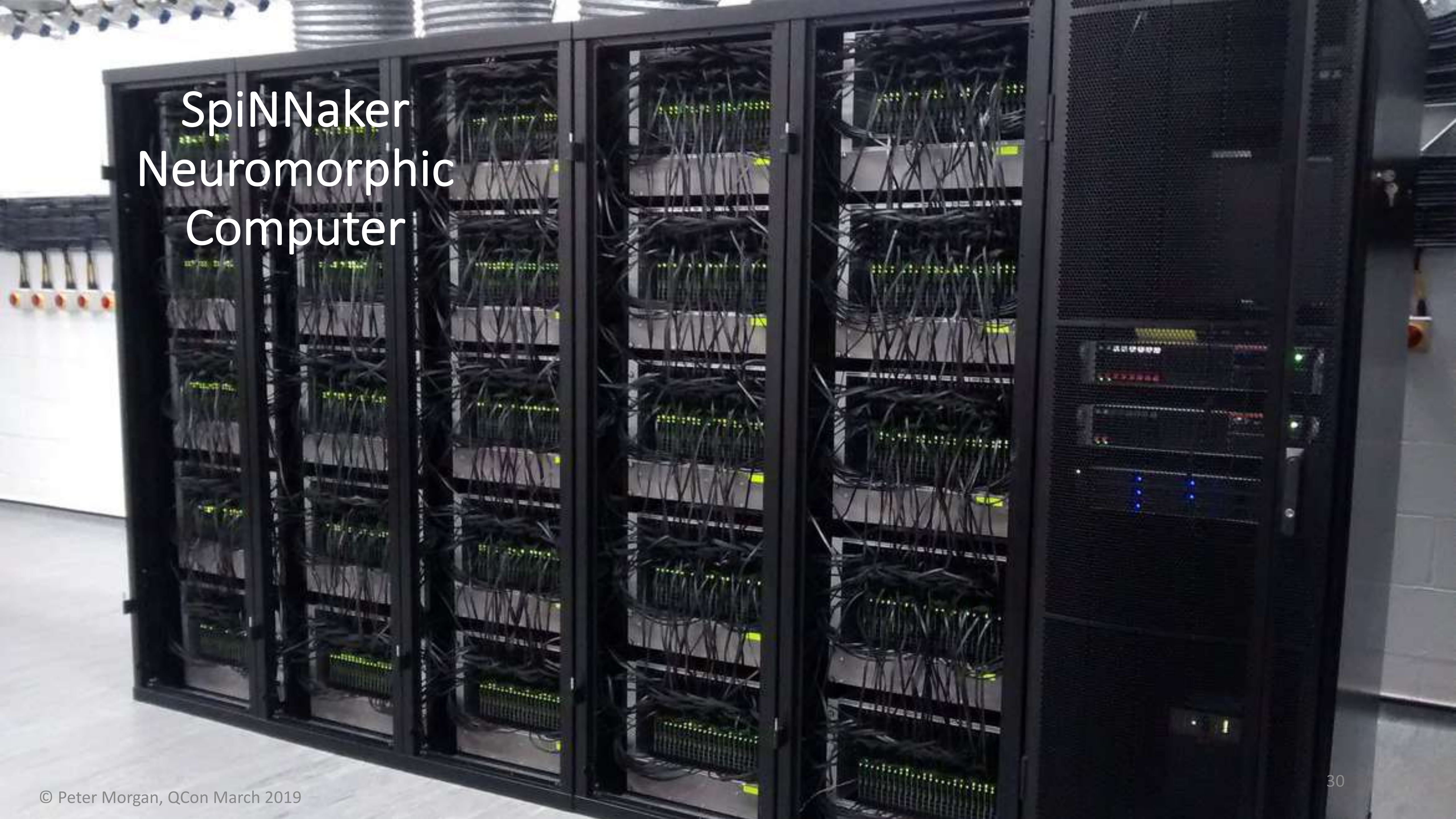
What are we missing?

- Brain computation is of order of PetaFlops, 1.5kg, 30W
- Clearly, shear digital horsepower (von Neumann architecture) is NOT going to get us to generally intelligent systems, or we would have them by now (have ExaFlops)
- We already know the brain (biology) uses a different architecture - memory and compute are combined (memristors)
- Perhaps we should be looking at the brain for inspiration - bioplausible architectures

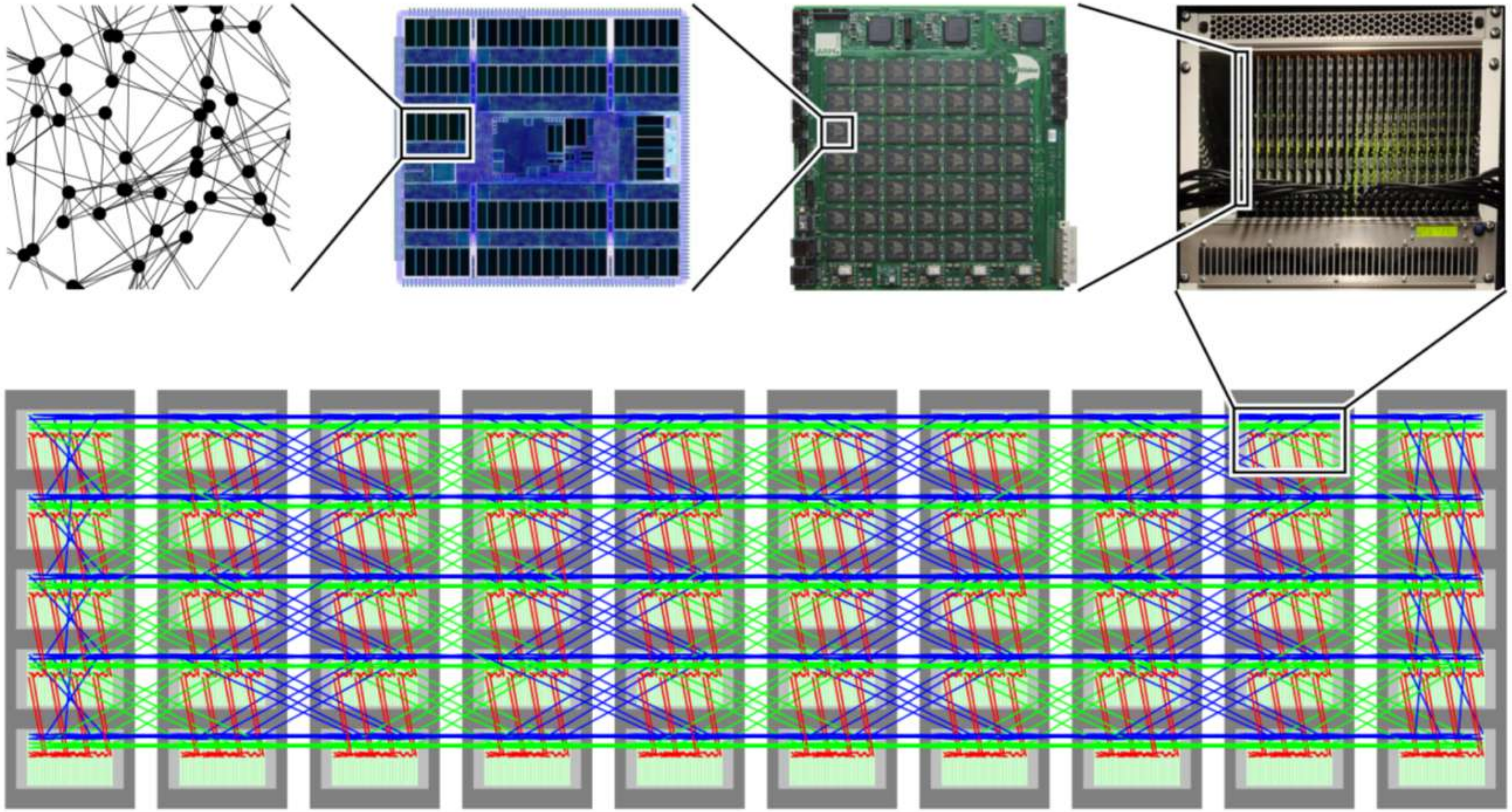
Introducing Neuromorphic Computing

- Biologically inspired
- First proposed Carver Mead, Caltech, 1980's
- Uses analogue signals – spiking neural networks (SNN)
  - SpiNNaker, BrainScaleS, TrueNorth, Intel Loihi
  - Startups - Knowm, Spaun, …
- Up to 1 million cores, 1 billion "neurons" (mouse)
- Need to scale 100X → human brain
- Relatively low power – 1,000X less than digital
- **SpiNNaker is available today in the cloud – try it out!**

SpiNNaker Neuromorphic Computer

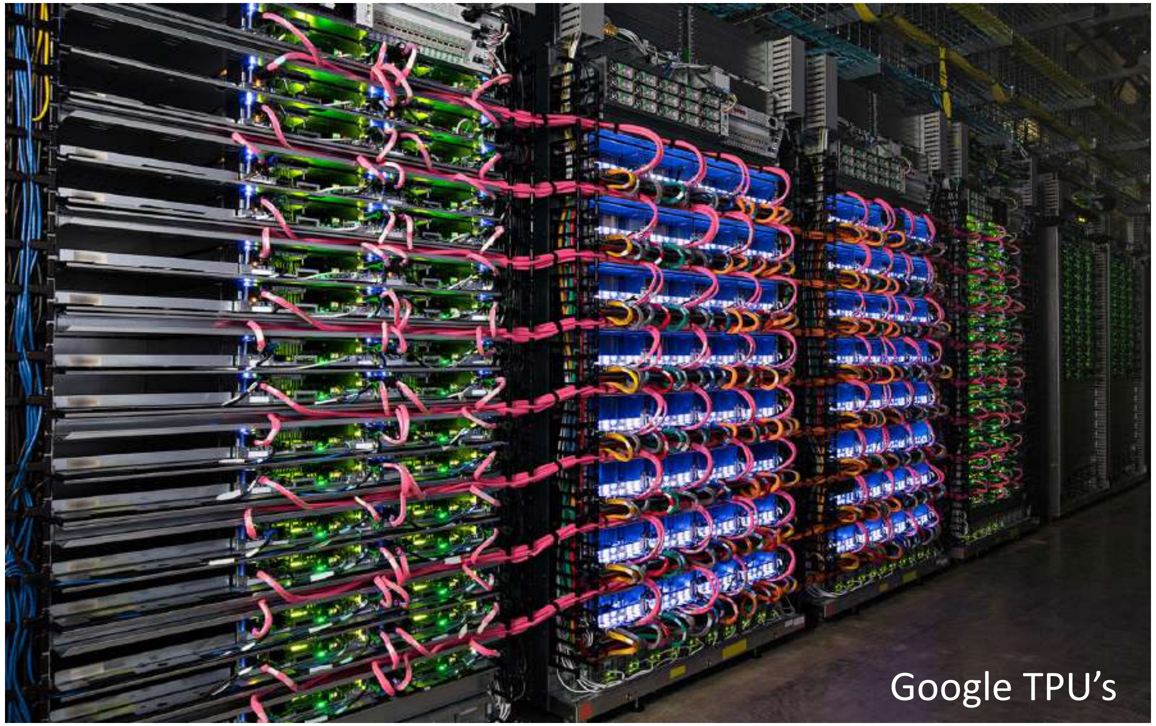


Scaling to a billion neurons



5 racks per cabinet, 10 cabinets.

11

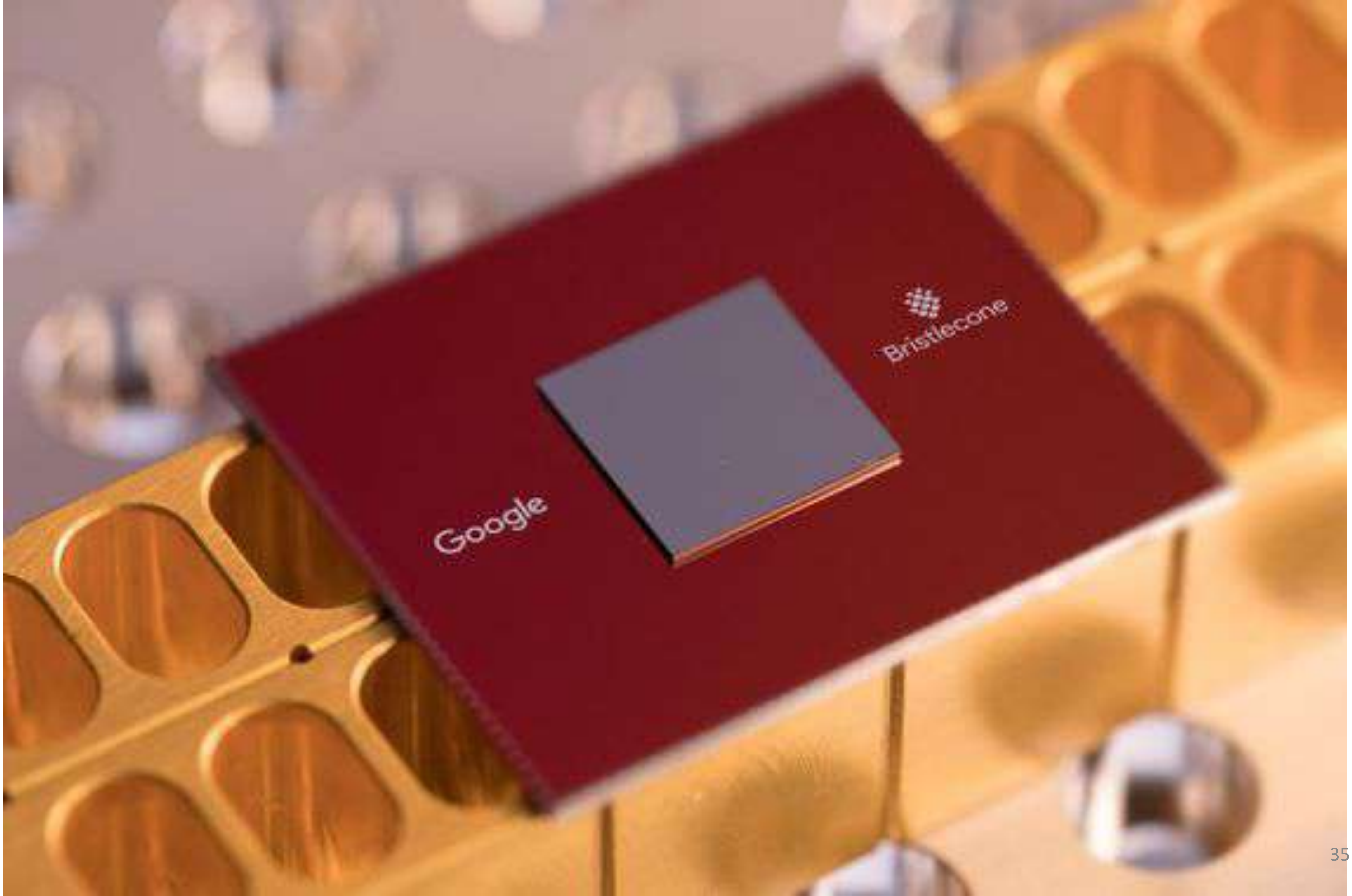


Analog v Digital

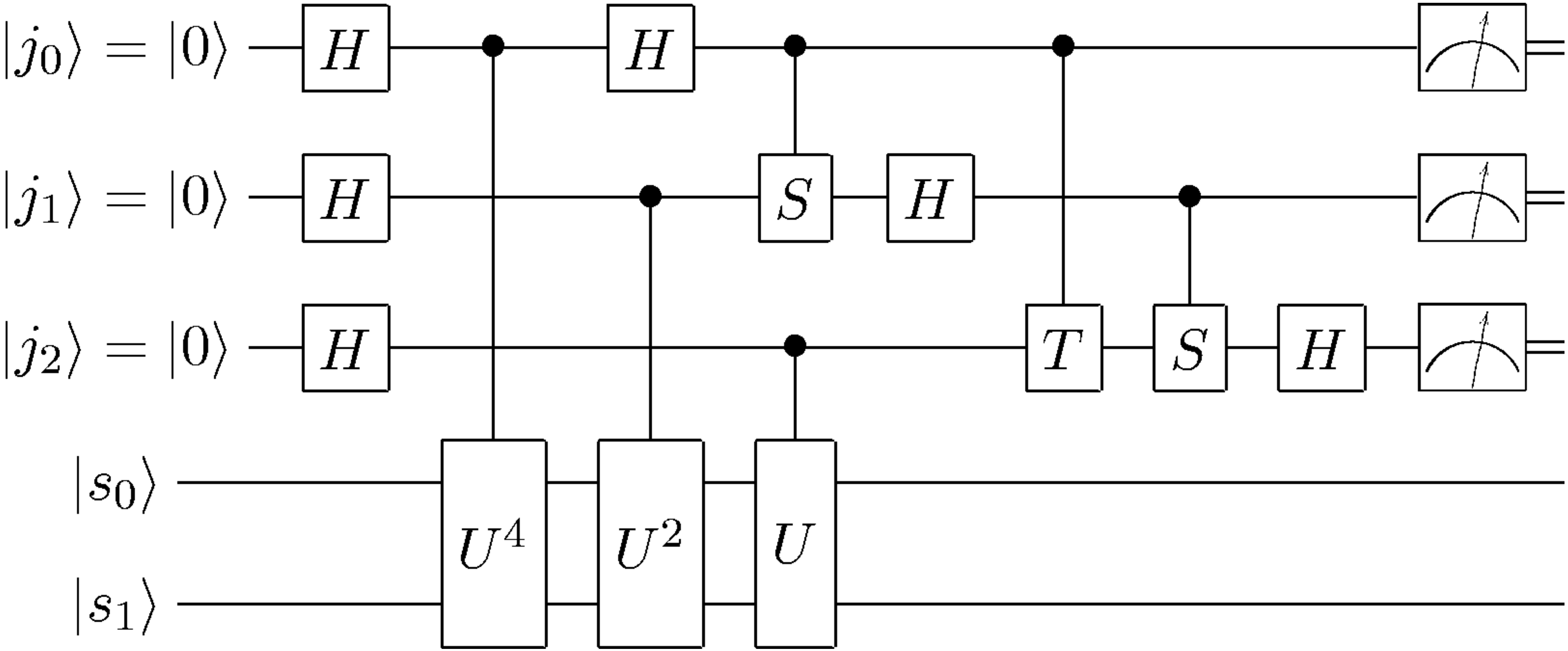
What about Quantum Computing?

- First proposed by Richard Feynman, Caltech, 1980’s
- Qubits – spin 1, 0, **and superposition states**
- Nature is fundamentally probabilistic at atomic scale
- Several approaches - superconductors, trapped ions, photonic, topological
- Several initiatives
  - IBM, D-Wave, Rigetti, Google, Intel, Microsoft, ...
- Applications – optimization, drug discovery, machine learning, cryptography
- Does Nature use quantum computation?
- See my talk tomorrow @ 3pm.

IBM 50 Qubit Quantum Computer


Google
Bristlecone

Quantum Logic Gates

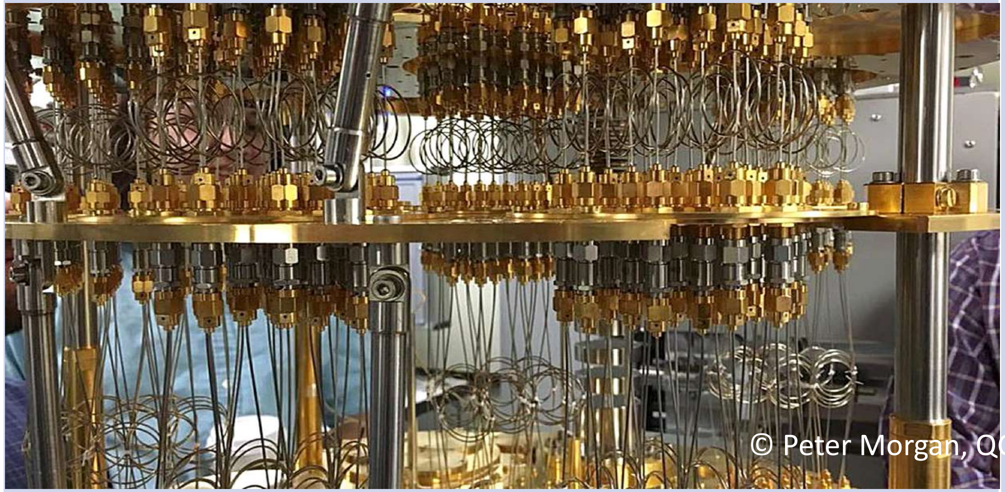


Summary – Four classes of physical computation systems

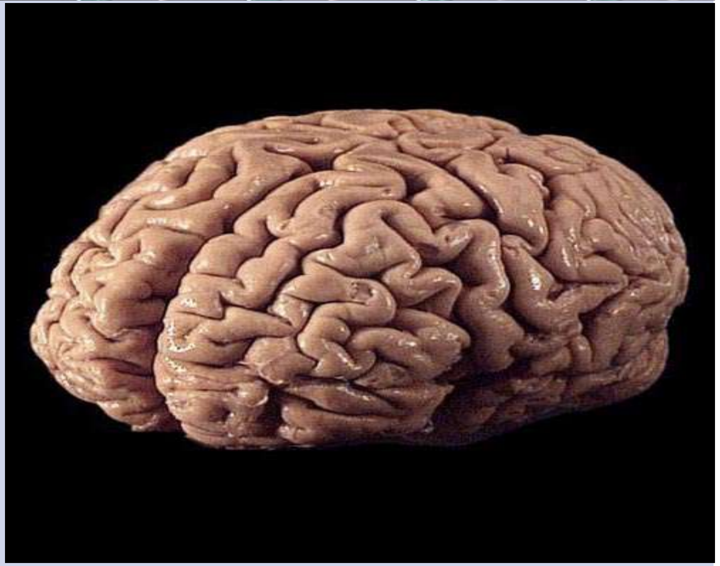
Digital

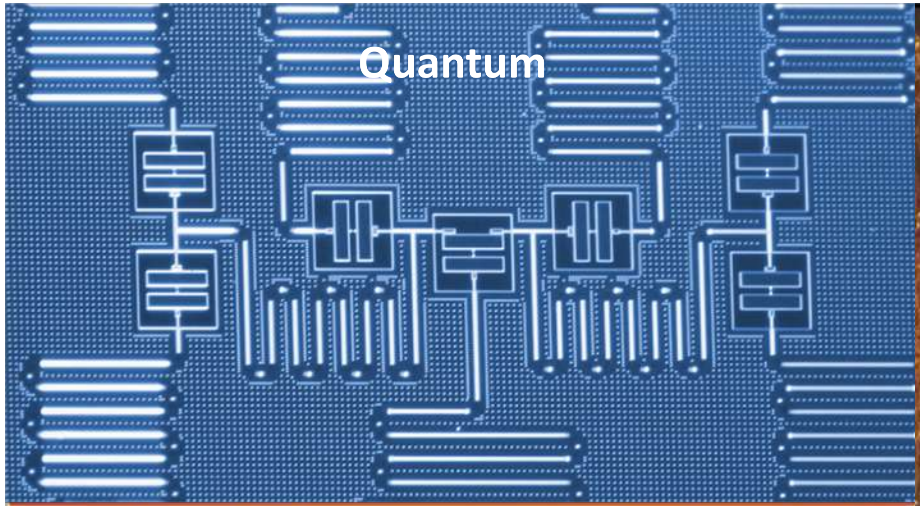
Neuromorphic

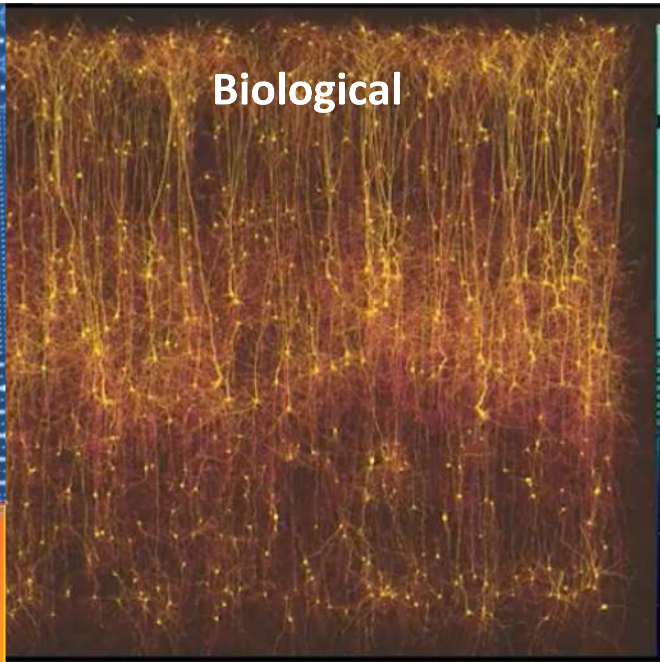
Quantum

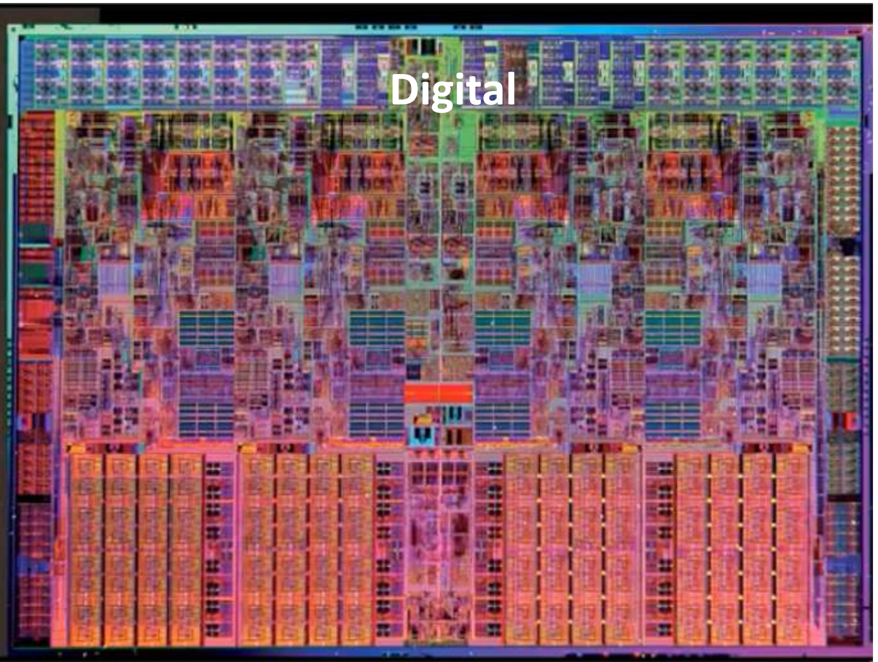


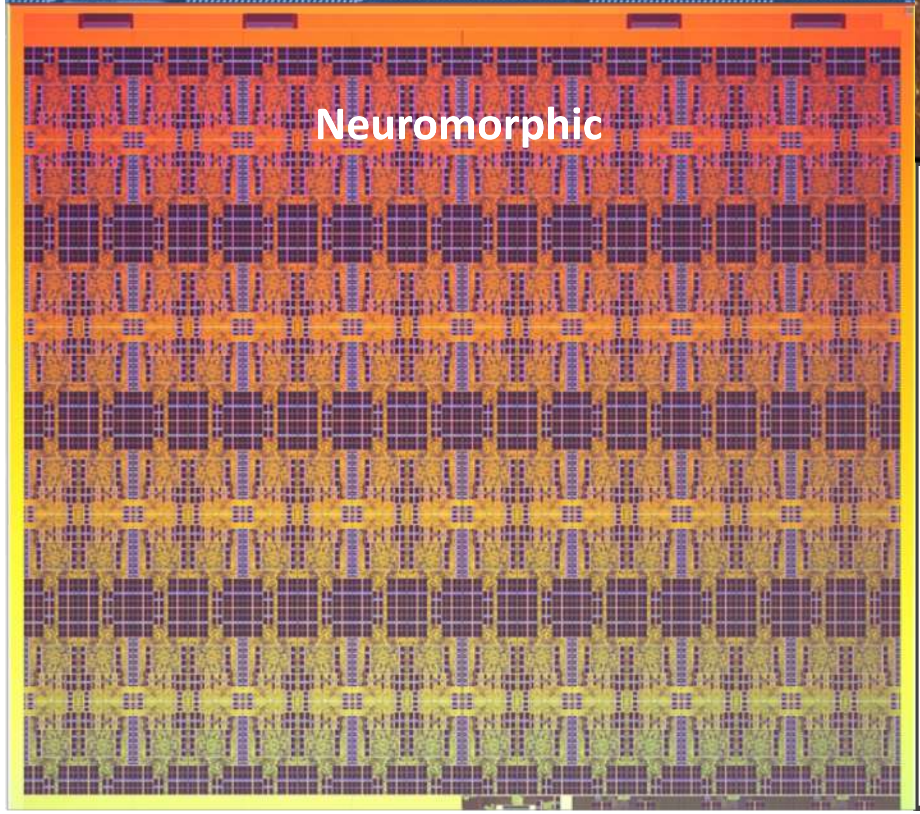
Biological





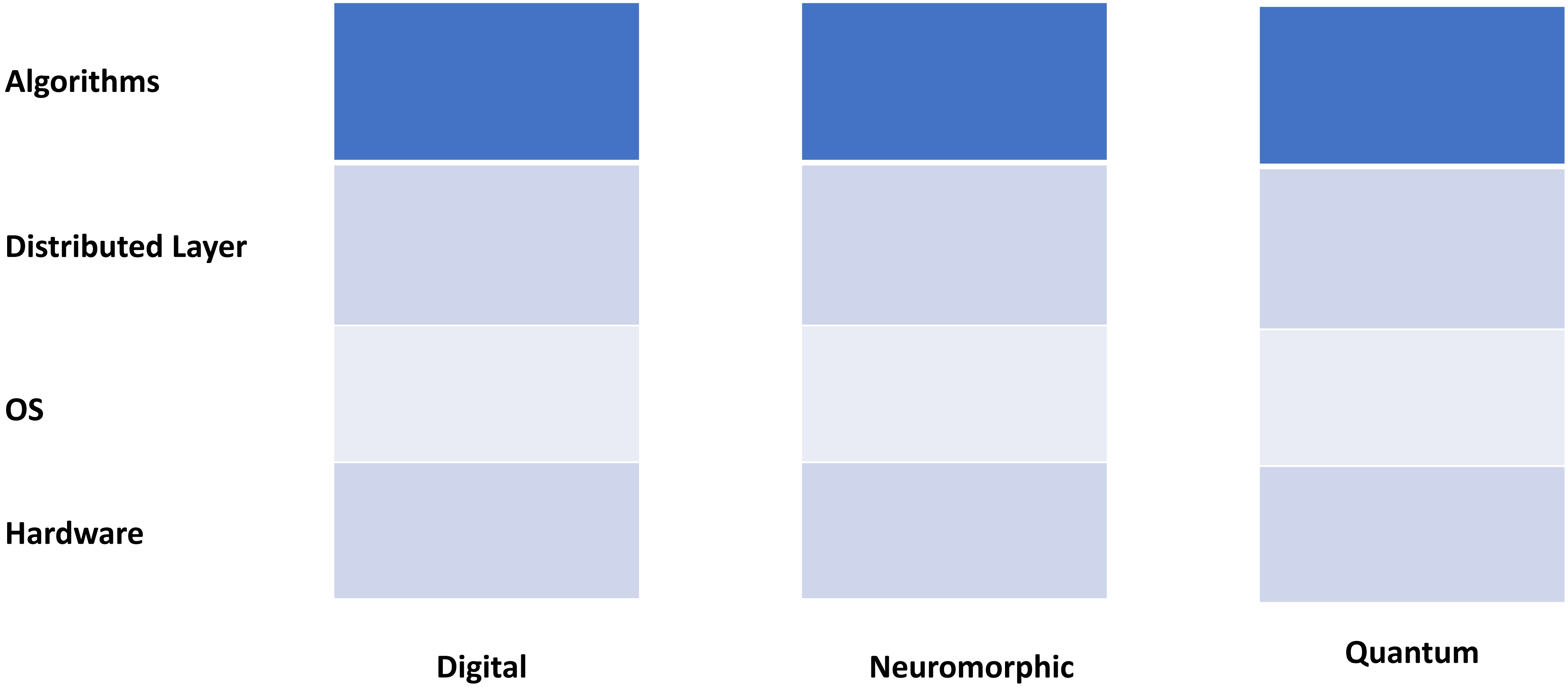


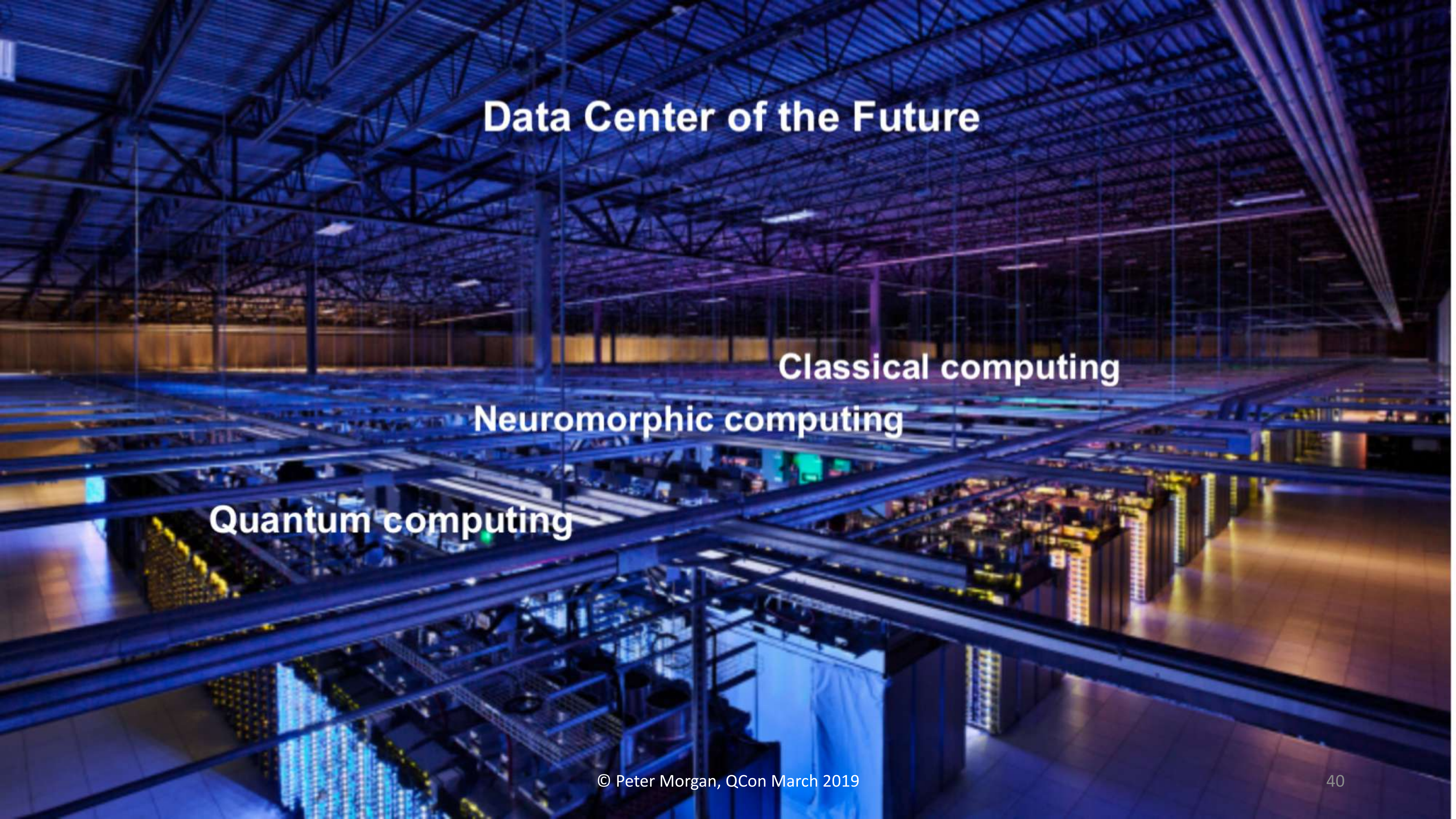




Four Types of Processors
All information processing systems

Three Non-biological Stacks




Data Center of the Future
Classical computing
Neuromorphic computing
Quantum computing

Outline

- What is Intelligence?
- Physical Systems
  - Biological
  - Non-biological
- **Deep Learning**
  - Recap
- AGI
  - Overview
  - Active Inference
  - Building AGI
- Conclusions

Early papers

NATURE VOL. 323 9 OCTOBER 1986 — LETTERS TO NATURE — 533

Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams*

* Institute for Cognitive Science, C-015, University of California, San Diego, La Jolla, California 92093, USA
† Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Philadelphia 15213, USA

Psychological Review
Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN[1]

F. ROSENBLATT

Cornell Aeronautical Laboratory

Psychological Review
1981, Vol. 88, No. 2, 135–170

Copyright 1981 by the American Psychological Association, Inc.
0033-295X/81/8802-0135$00.75

Toward a Modern Theory of Adaptive Networks: Expectation and Prediction

Richard S. Sutton and Andrew G. Barto
Computer and Information Science Department
Unive

COGNITIVE SCIENCE 9, 147–169 (1985)

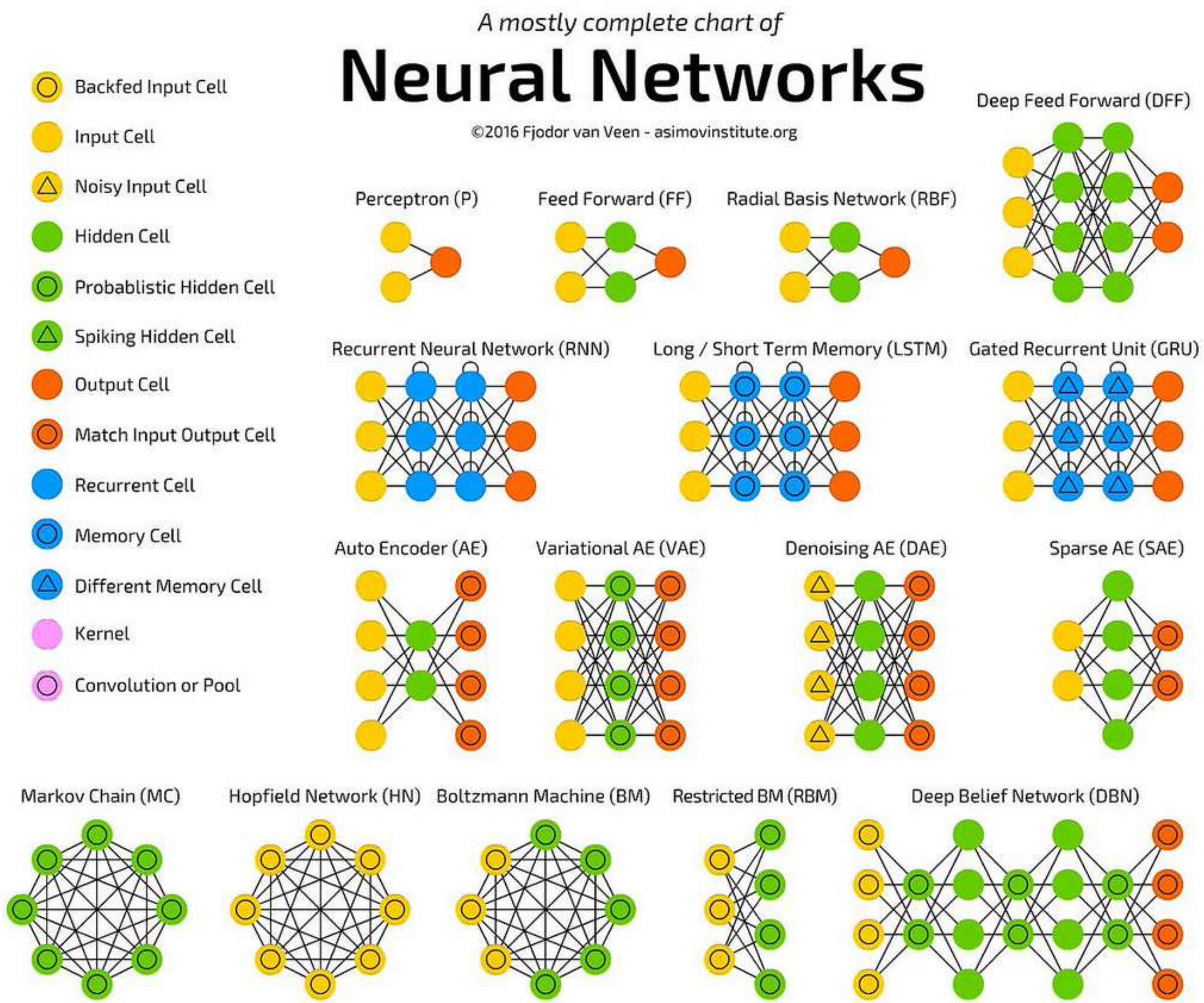
A Learning Algorithm for Boltzmann Machines*

DAVID H. ACKLEY
GEOFFREY E. HINTON
Computer Science Department
Carnegie-Mellon University
TERRENCE J. SEJNOWSKI
Biophysics Department
The Johns Hopkins University

COGNITIVE SCIENCE 14, 179–211 (1990)

Finding Structure in Time

JEFFREY L. ELMAN
University of California, San Diego


A mostly complete chart of
Neural Networks
©2016 Fjodor van Veen - asimovinstitute.org
Backfed Input Cell
Input Cell
Noisy Input Cell
Hidden Cell
Probablistic Hidden Cell
Spiking Hidden Cell
Output Cell
Match Input Output Cell
Recurrent Cell
Memory Cell
Different Memory Cell
Kernel
Convolution or Pool
Perceptron (P)
Feed Forward (FF)
Radial Basis Network (RBF)
Deep Feed Forward (DFF)
Recurrent Neural Network (RNN)
Long / Short Term Memory (LSTM)
Gated Recurrent Unit (GRU)
Auto Encoder (AE)
Variational AE (VAE)
Denoising AE (DAE)
Sparse AE (SAE)
Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)

Deep Learning Performance

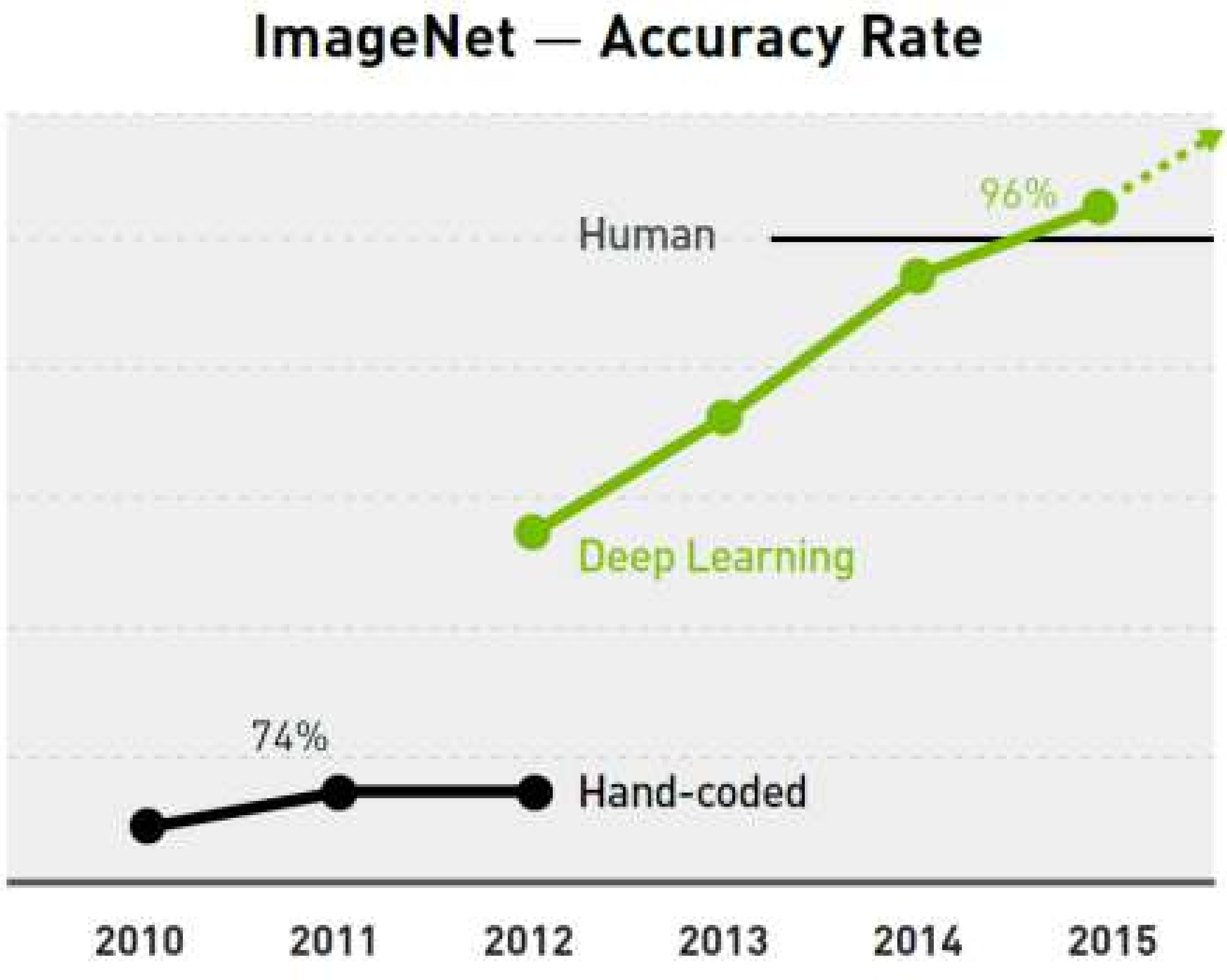
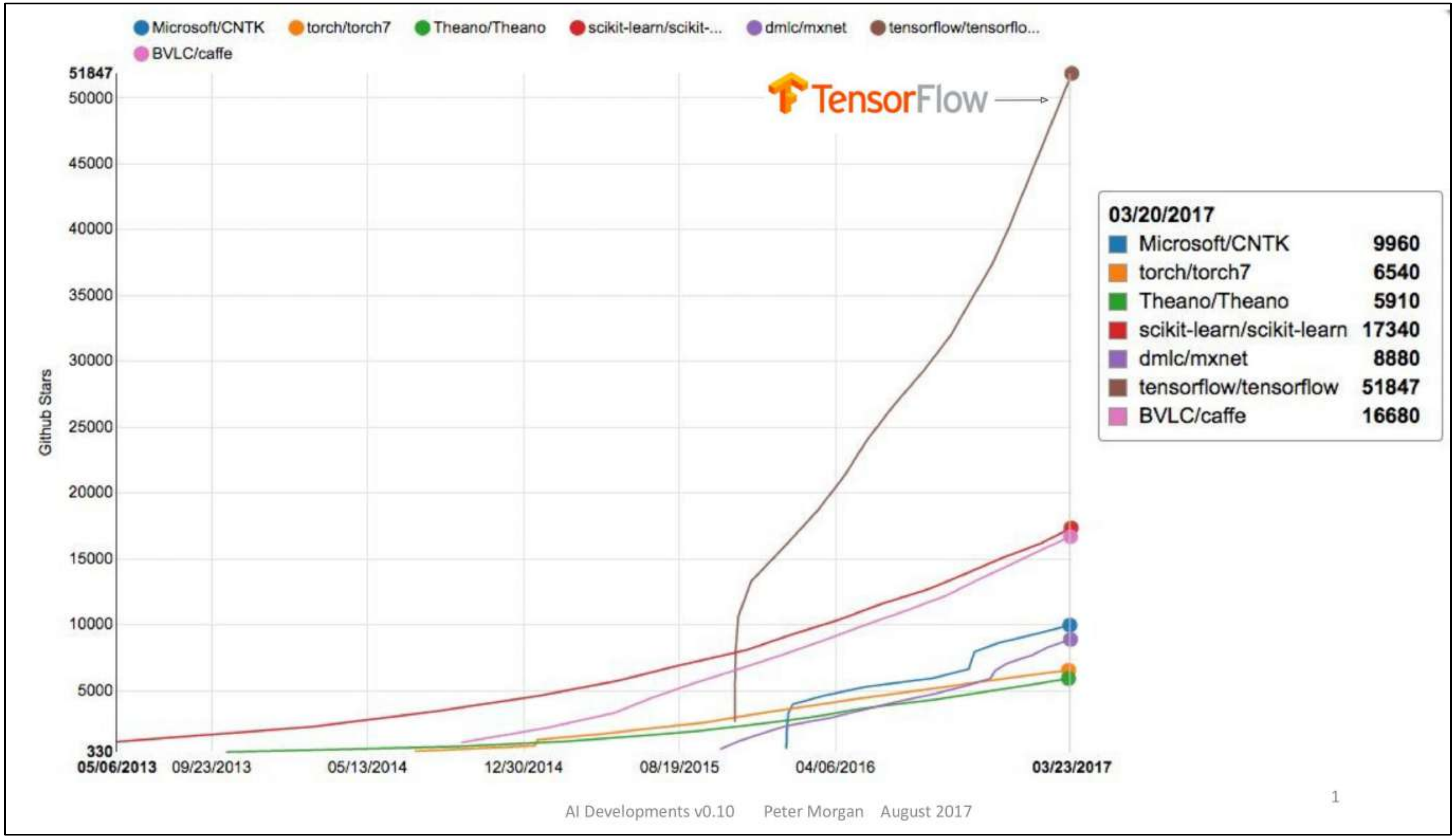


Image classification

Framework Popularity



TensorFlow 2.0 is released this week!

The fourth industrial revolution will be (is) open source

- ML Frameworks – open source (e.g., TensorFlow)
- Operating systems – open source (Linux)
- Hardware – open source
  - OCP (Open Compute Project) & RISC-V
- Data sets – open source (Internet)
- Research – open source (see arXiv, etc.)
- Is tremendously accelerating progress

Outline

- What is Intelligence?
- Physical Systems
  - Biological
  - Non-biological
- Deep Learning
  - Recap
- **AGI**
  - Overview
  - Active Inference
  - Building AGI
- Conclusions

General Theories of Intelligence

- What do we need?
- Different approaches
- Active Inference
- Building AGI

AGI = Artificial General Intelligence

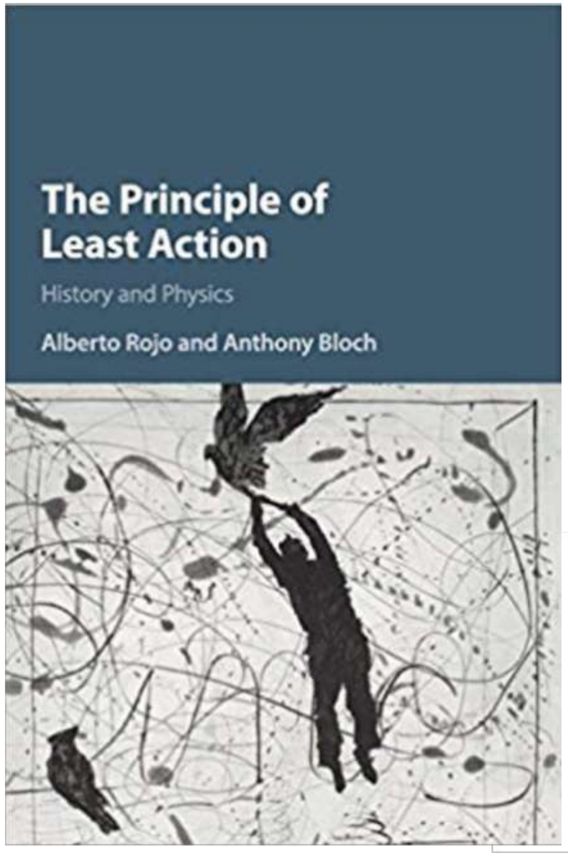
The Physics Approach

The brain is a physical system so what are the fundamental physical principles?

- Newtonian mechanics – three laws
- Special relativity – invariance of laws under a Lorentz transformation
- GR – Principle of Equivalence
- Electromagnetism – Maxwell's equations
- Thermodynamics – three laws
- Quantum mechanics – uncertainty principle
- Relativistic QM – Dirac equation
- Dark energy/dark matter – we don't know yet

All of the above = Principle of Least Action

The Principle of Least Action
(All of physics can be derived from this)



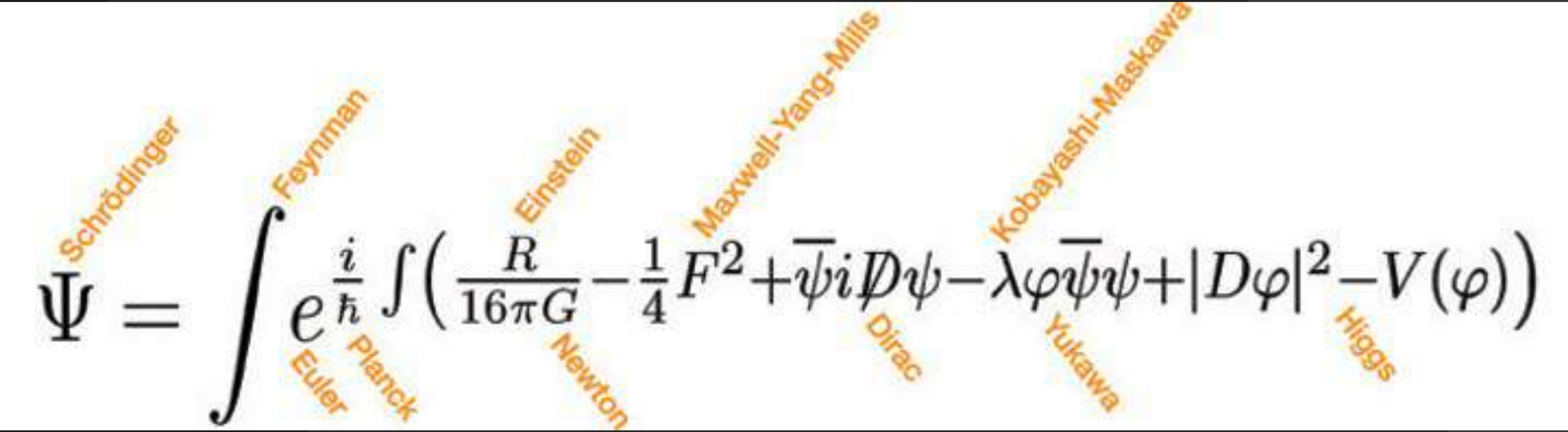
Cambridge University Press, 2018

$$\delta S = 0,$$

$$S[\mathbf{q}, t_1, t_2] = \int_{t_1}^{t_2} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)dt$$

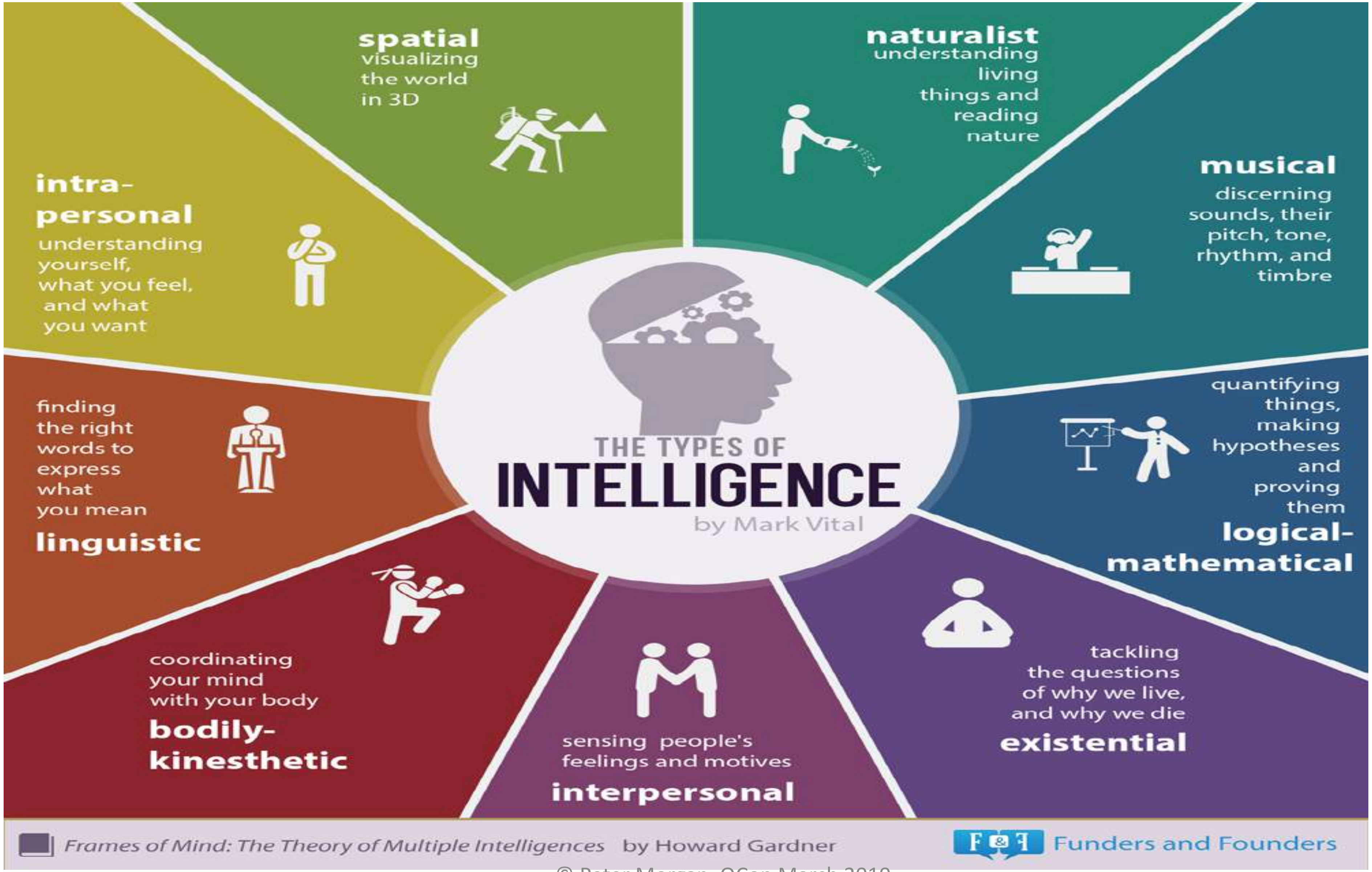
Please don't worry about the math in this section – try to focus on the concepts

All known physics – Field theoretic



Pinnacle of human achievement?

Applied to Intelligence



We need a system that can model the (whole) world

Intelligence is not just about ***pattern recognition***

It is about ***modeling*** the world

- **Explaining** and **understanding** what we see
- **Imagining** things we could see but haven't yet
- **Problem solving** and **planning actions** to make these things real
- **Building new models** as we learn more about the world

Theoretical Approaches to AGI

- Helmholtz and others (Statistical physics, late 1800's)
- **Friston – Active Inference**
- Tishby – Information bottleneck
- Bialek – Biophysics
- Hutter – AIXI
- Schmidhuber – Godel Machine
- All of the above have been worked on for thirty+ years
- Many others

Active Inference

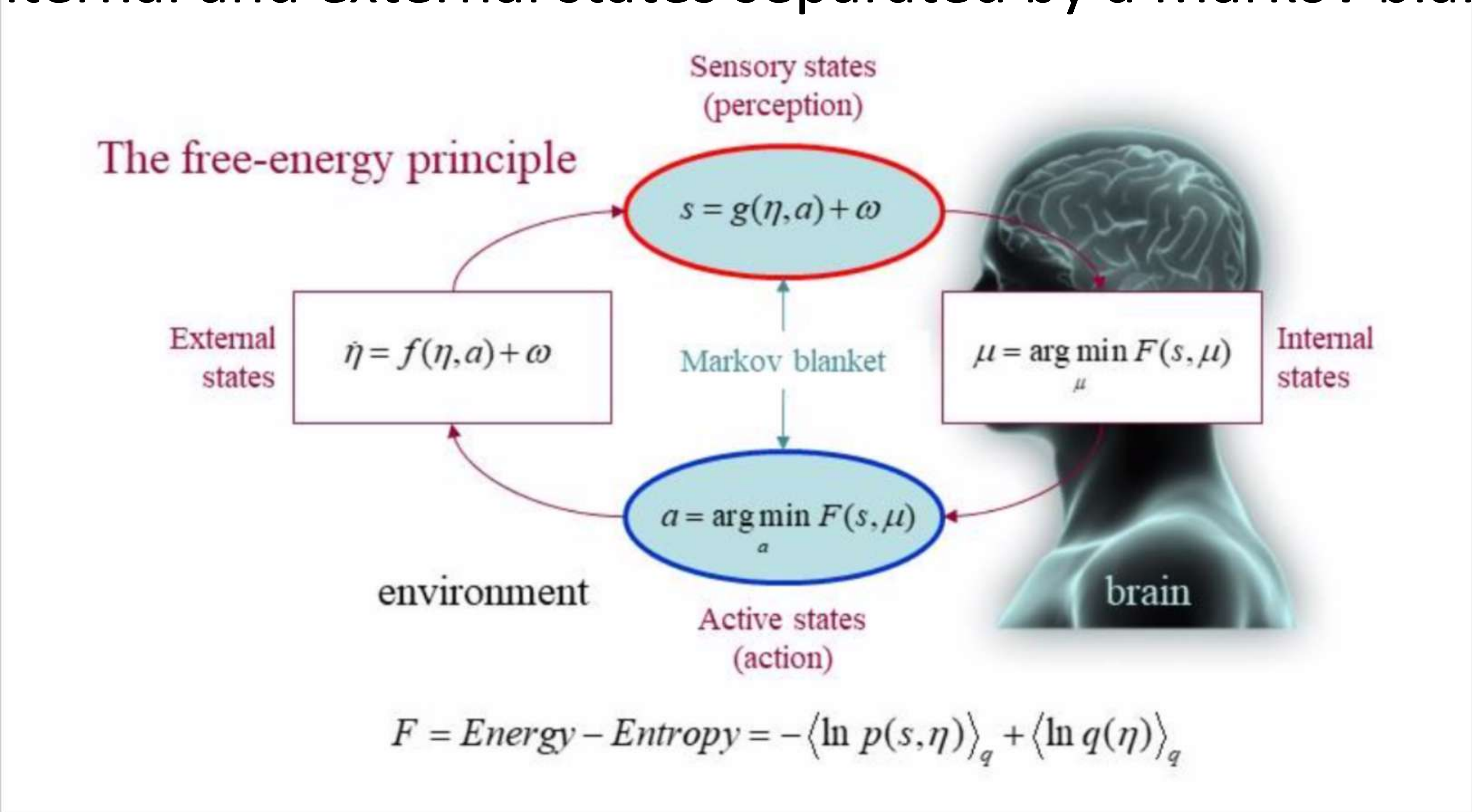
- Professor Karl Friston, UCL Neuroscience
- Based on physics and information theory
- Uses the Free Energy Principle
  - Systems act to minimize their expected free energy
  - Reduce uncertainty (prediction error)
- It encompasses all interactions and dynamics
  - Completely general
  - Applies over all time and distance scales


Karl Friston
Professor of Neurology
University College London

Active Inference Summary

- Biological agents resist the **second law of thermodynamics**
- They must minimize their average surprise **(entropy)**
- They minimize surprise by suppressing prediction error **(free-energy)**
- Prediction error can be reduced by changing predictions **(perception)**
- Prediction error can be reduced by changing sensations **(action)**
- Perception recurrent message passing in brain - **optimizes predictions**
- Action makes predictions come true (and **minimizes surprise)**

Internal and external states separated by a Markov blanket



Active Inference - Information theoretic approach, uses generalised free energy

$$Q(\pi) = \arg\min \sum_\tau E_Q[F(\pi,\tau)] + D[Q(\pi) \| P(\pi)]$$

Policy selection

$$Q(s_\tau | \pi) = \arg\min \sum_\tau F(\pi,\tau)$$

Perceptual inference

$$F(\pi,\tau) = E_Q[\underbrace{\ln Q(s_\tau | \pi)}_{entropy} - \underbrace{\ln P(o_\tau, s_\tau | \pi)}_{energy}]$$

$$Q(o_\tau | s_\tau) = \begin{cases} P(o_\tau | s_\tau) & : \tau > t \\ \delta(o_\tau) & : \tau \le t \end{cases}$$

$$P(s_\tau | \pi) = \begin{cases} P(s_\tau) & : \tau > t \\ P(s_\tau | \pi) & : \tau \le t \end{cases}$$

$$\ln P(\pi) = \sum_\tau -G(\pi,\tau)$$

$$G(\pi,\tau) = E_Q[\underbrace{\ln Q(s_\tau | \pi)}_{entropy} - \underbrace{\ln P(o_\tau, s_\tau | \pi)}_{energy}]$$

$$\begin{aligned} G(\pi,\tau) &= E_Q[F(\pi,\tau)] \\ &= E_{Q(o_\tau, s_\tau | \pi)}[\underbrace{\ln Q(o_\tau | \pi) Q(s_\tau | \pi)}_{entropy} - \underbrace{\ln P(o_\tau, s_\tau)}_{energy}] \\ &= E_{Q(o_\tau, s_\tau | \pi)} \underbrace{D[Q(s_\tau | \pi) \| P(s_\tau | \pi)]}_{expected\ cost} - \underbrace{D[Q(o_\tau, s_\tau) \| Q(s_\tau | \pi) Q(o_\tau | \pi)]}_{epistemic\ value\ (mutual\ information)} \end{aligned}$$

Generalised free energy – with some care

Warning: the maths gets ugly

A

Variational updates

Perception (state-estimation)

$$\mathbf{s}_\tau^{n,\pi} = \sigma(\hat{\mathbf{s}}_\tau^{n,\pi} - \tfrac{1}{4}\nabla F)$$

$$\nabla F(\pi,\tau,n) = \hat{\mathbf{s}}_\tau^{n,\pi} - \hat{\mathbf{B}}_{\tau-1}^{n,\pi}\mathbf{s}_{\tau-1}^{n,\pi} - \hat{\mathbf{B}}_\tau^{n,\pi}\cdot\mathbf{s}_{\tau+1}^{n,\pi} - \sum\nolimits_m (\hat{\mathbf{A}}^m \cdot \mathbf{s}_\tau^{/n,\pi})\cdot o_\tau^m$$

Policy selection (and evaluation)

$$\boldsymbol{\pi} = \sigma(-\mathbf{F} - \boldsymbol{\gamma}\cdot\mathbf{G})$$

$$\mathbf{F}_\pi = \sum\nolimits_{n,\tau} \nabla F(\pi,\tau,n)\cdot\mathbf{s}_\tau^{n,\pi}$$

$$\mathbf{G}_\pi = \sum\nolimits_{m,\tau} (\mathbf{o}_\tau^{m,\pi}\cdot(\hat{\mathbf{o}}_\tau^{m,\pi} - \hat{\mathbf{C}}_\tau^m) + \mathbf{H}^m\cdot\mathbf{s}_\tau^{\bullet,\pi})$$

$$\mathbf{o}_\tau^{m,\pi} = \mathbf{A}^m\cdot\mathbf{s}_\tau^{\bullet,\pi}$$

Precision (softmax parameter)

$$\boldsymbol{\beta} = \boldsymbol{\beta} - \tfrac{1}{4}\nabla F$$

$$\nabla F = (\boldsymbol{\beta} - \beta) - (\boldsymbol{\pi} - \boldsymbol{\pi}_0)\cdot\mathbf{G}$$

$$\boldsymbol{\pi}_0 = \sigma(-\boldsymbol{\gamma}\cdot\mathbf{G})$$

Action selection (and model averaging)

$$a_t = \min\nolimits_a \sum\nolimits_m \mathbf{o}_{t+1}^m\cdot(\hat{\mathbf{o}}_{t+1}^m - \hat{\mathbf{o}}_{t+1}^{m,a})$$

$$\mathbf{s}_{t+1}^n = \sum\nolimits_\pi \boldsymbol{\pi}_\pi\cdot\mathbf{s}_{t+1}^{n,\pi}$$

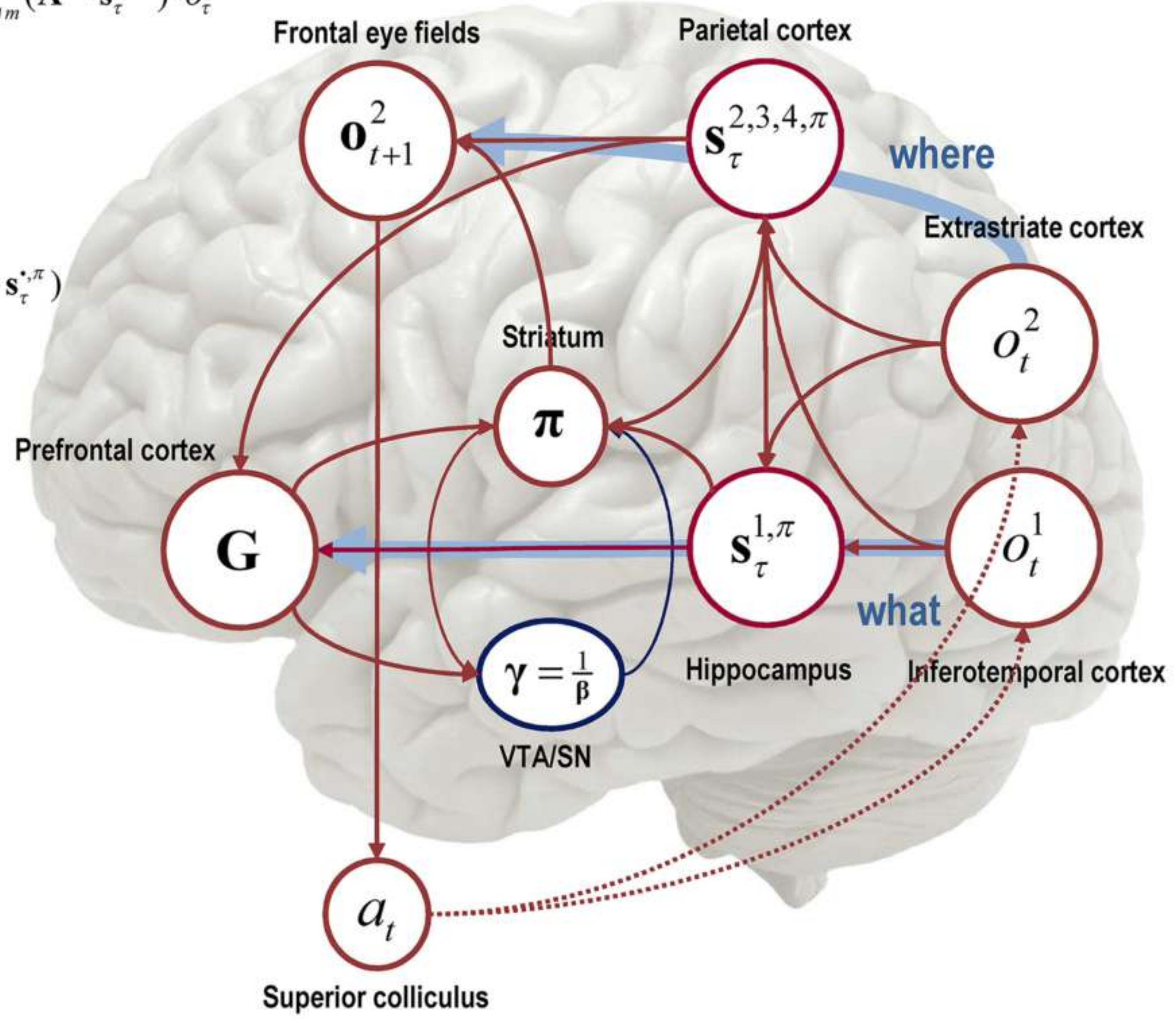
$$\mathbf{s}_{t+1}^{n,a} = \mathbf{B}^n(a)\mathbf{s}_t^n$$

$$\mathbf{o}_{t+1}^m = \mathbf{A}^m\cdot\mathbf{s}_{t+1}^\bullet$$

$$\mathbf{o}_{t+1}^{m,a} = \mathbf{A}^m\cdot\mathbf{s}_{t+1}^{\bullet,a}$$

B

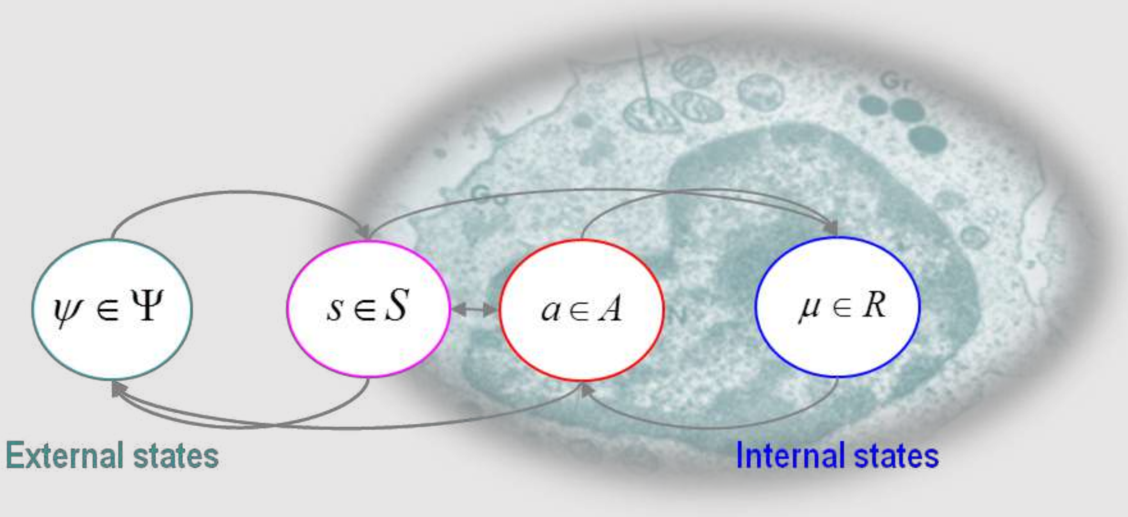
Functional anatomy



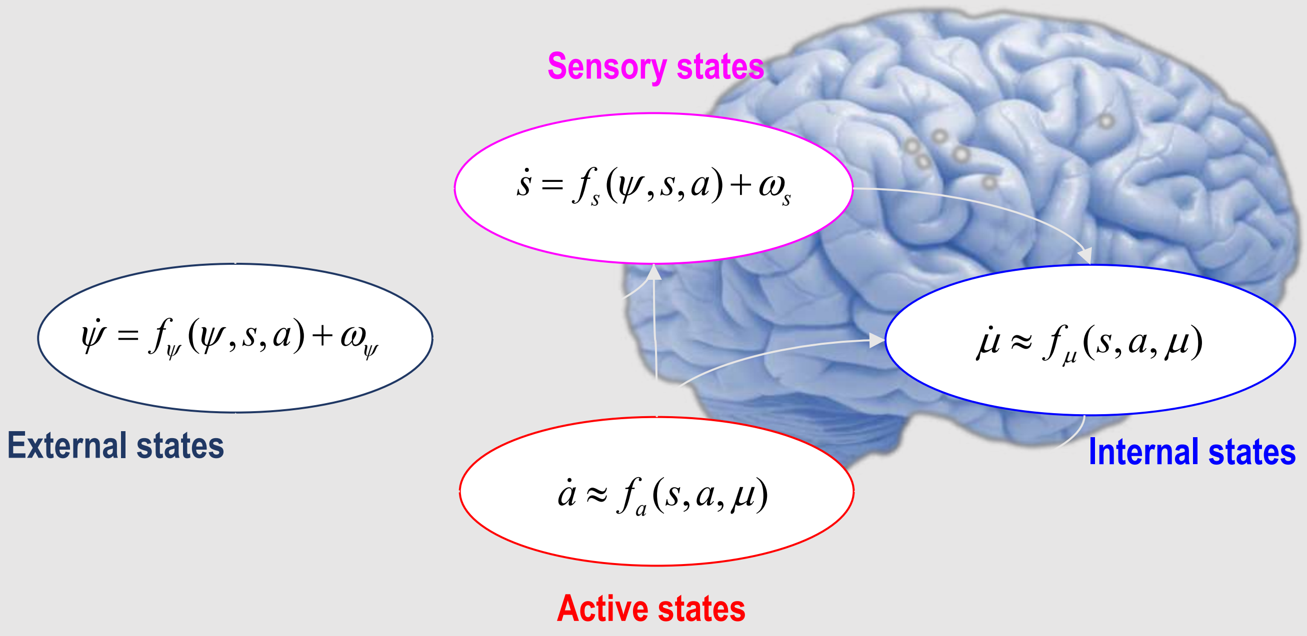
Active Inference

We have code

From cells to brains



Cell



Brain

Can we build general intelligence?

- We have
  - Candidate theories – active inference ✓
  - Algorithms & software ✓
  - Hardware – ASIC, neuromorphic ✓
  - Data sets ✓
- Need to build a complete framework with libraries
  - A “TensorFlow for general intelligence”
  - **We need software engineers for this part ☺**
- Apollo Project of our time – “Fourth (Intelligence) Revolution”
  - Steam → electricity → digital → intelligence
  - Human Brain Project, Deepmind, BRAIN project, China, …
- ***Should*** we build AGI/ASI? – safety, ethics, singularity?
  - Topic for another talk(s) – see ethics track

Some AGI Projects

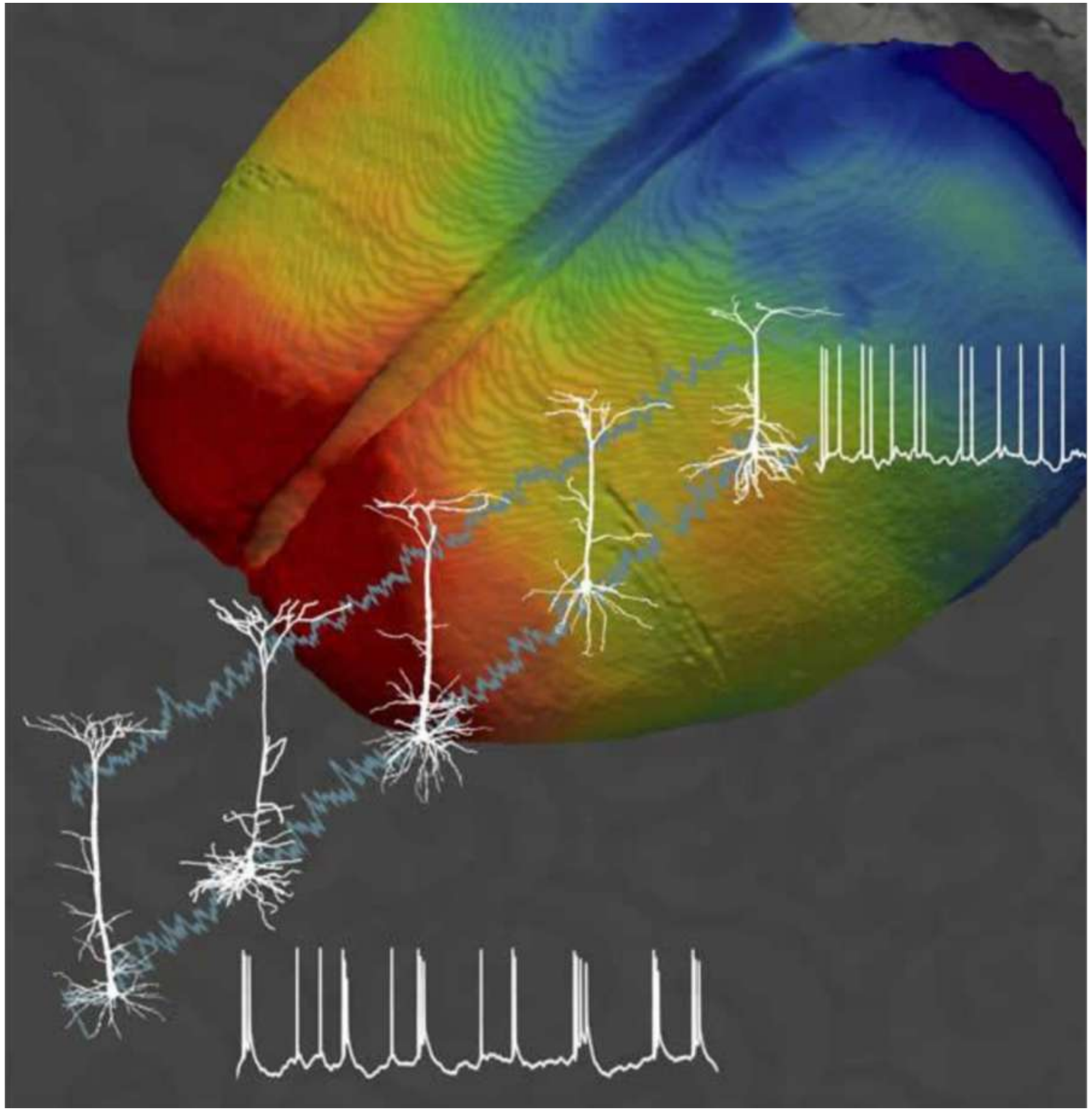
Deepmind – Demis Hassibis (UK)	OpenCog – Ben Goertzel (US)	Numenta – Jeff Hawkins (US)	Vicarious – Dileep George (US)	OpenAI – Sam Altman (US)
NNAIsense – Jurgen Schmidhuber (Swiss)	AGI Innovations – Peter Voss (US)	GoodAI – Marek Rosa (Czech)	Curious AI – (Finland)	Eurisko – Doug Lenat (US)
SOAR – CMU	ACT-R – CMU	Turing.AI (UK)	Sigma – Paul Rosenbloom, USC	Plus many, many more

Conclusions

- It is obvious to most that Deep Learning is lacking the foundations needed for a general theory of intelligence – it is based on statistics not physics
- Active inference ***is*** a theory of general intelligence
- Deep Learning research groups are now (finally) turning to biology for inspiration
- Bioplausible models are starting to appear
- Some groups are ***starting*** to look at active inference
- Real AGI systems in one year? five years? ten years?
- Still have to wait for hardware to mature
- Neuromorphic might be the platform that gets us there.

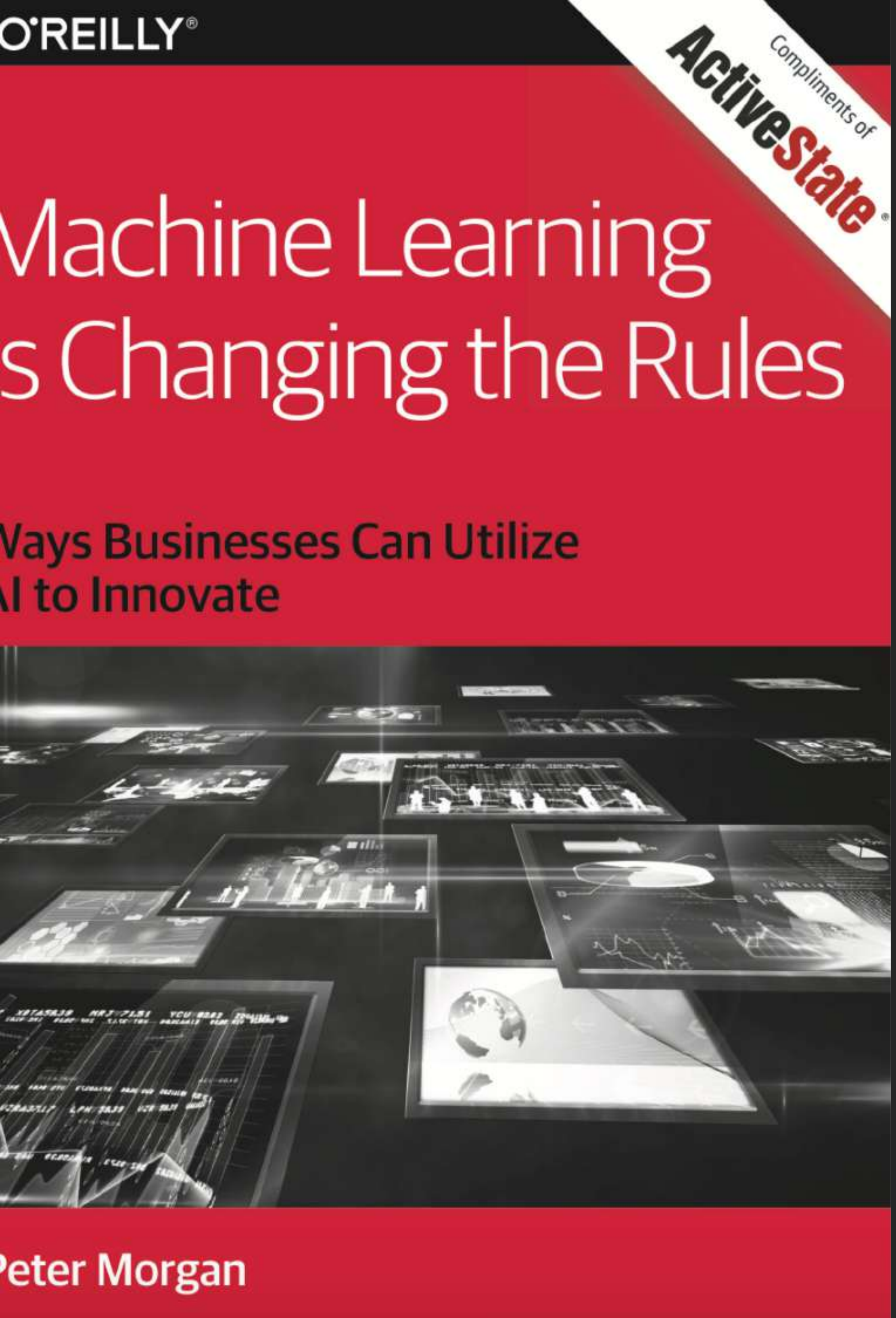
TL;DR

Using the physical principles of active inference, I believe we can build AGI systems over the coming years.



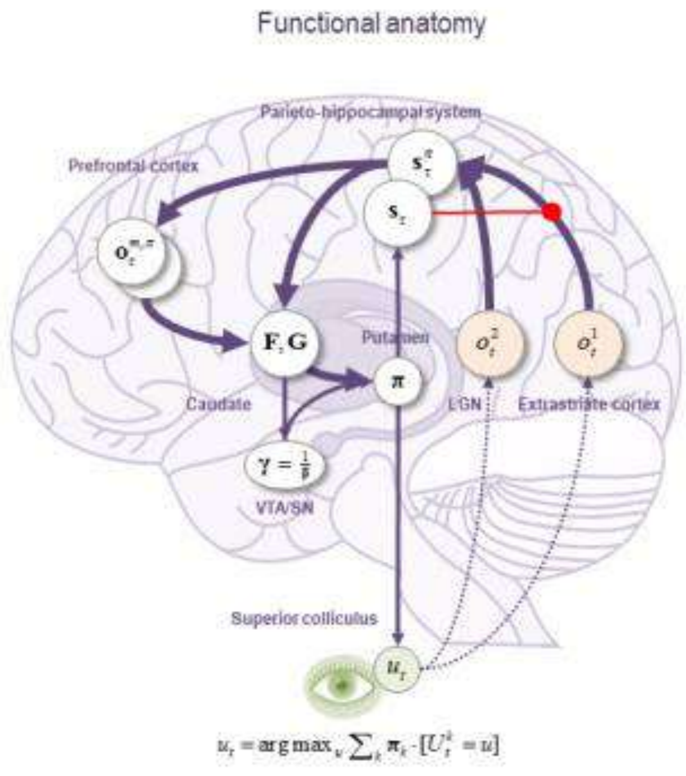
Final Word …

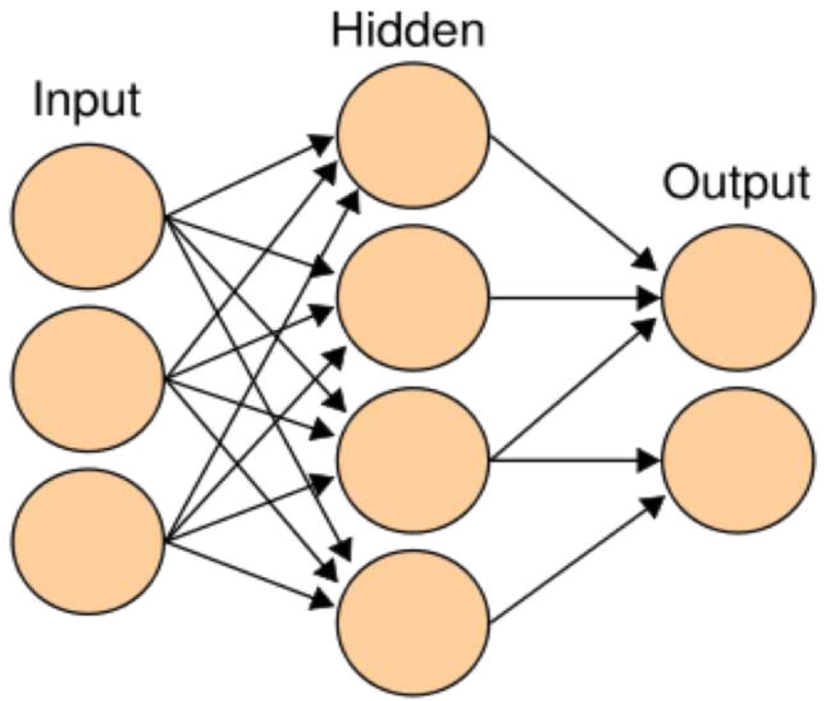

QUICK
INSIGHTS
Future of AI
"Assuming the computer industry can keep producing better hardware, I think 'business as usual' is going to take us a long way. Obviously, if we get big conceptual breakthroughs, it'll take us further. I think one of the big breakthroughs that's going to come is we're going to understand the brain."
Geoffrey Hinton (2016)
“Godfather of AI”



Report download
https://start.activestate.com/oreilly-ebook/

Company
www.turing-ai.co





AMA 10:35am in Guild Room