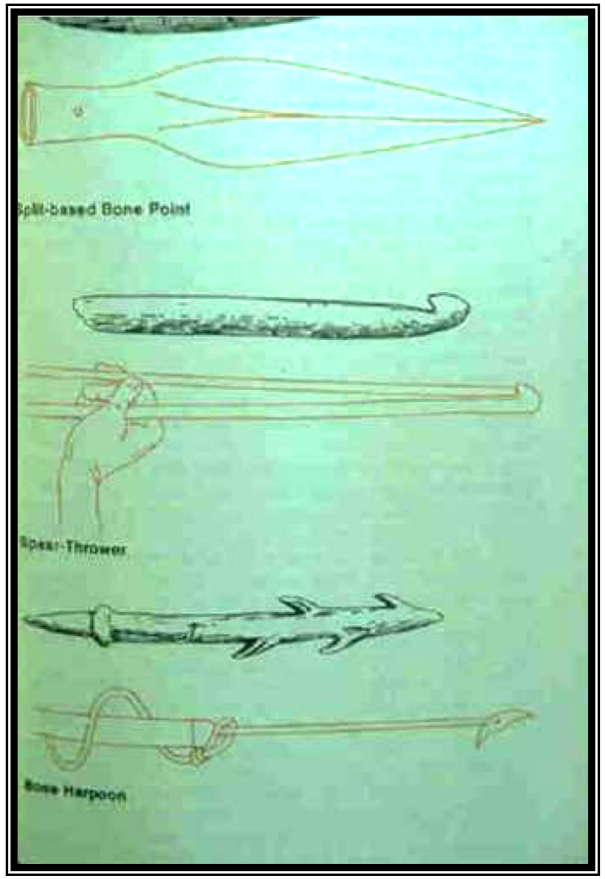
A Brief History of Neural Network Field

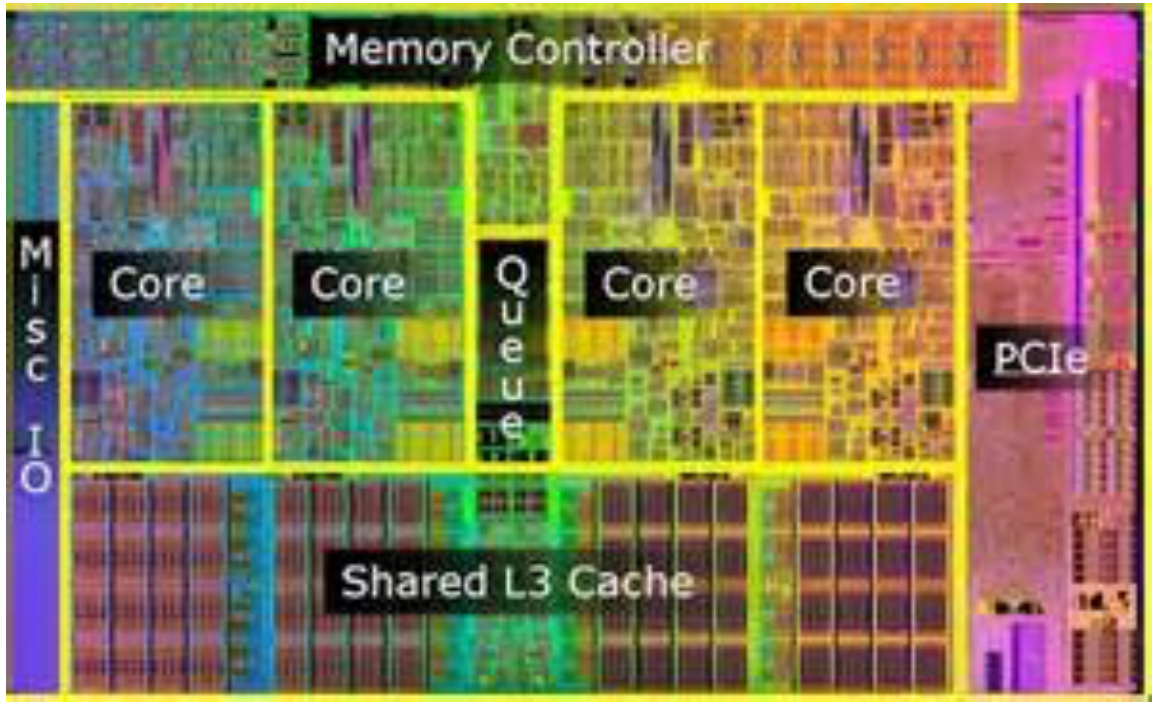
Jose C. Principe, Ph.D.
Distinguished Professor of ECE
Eckis Professor,
University of Florida, Gainesville
principe@cnel.ufl.edu

Disclaimer: All pictures where downloaded from publicly available websites

Humans and their Impressive Technologies

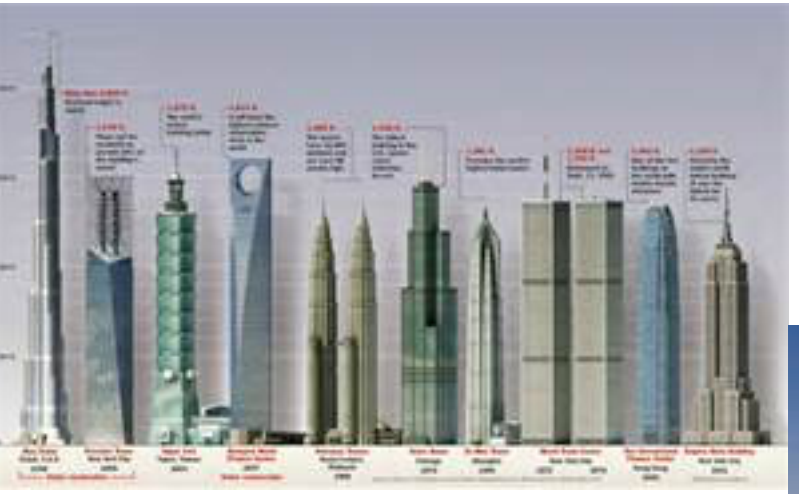
- Mostly a Master-Slave Relationship
- Trend is towards Brain Centric Technologies

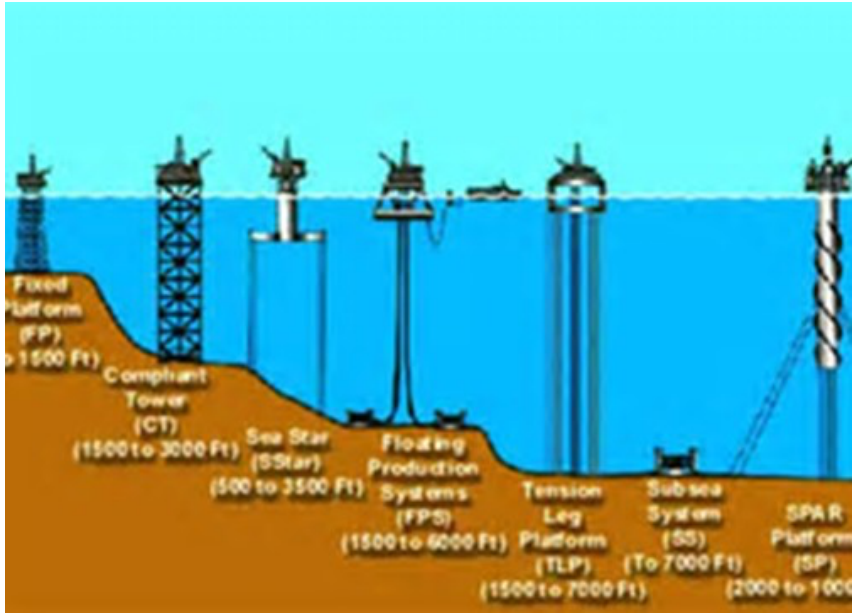




Intel i7 Core ~ 1.4 billion transistors

Our Technology World is Impressive





Burj Khalifa in Dubai, United Arab Emirates
world's tallest man-made structure is 829.8 m (2,722 ft) tall

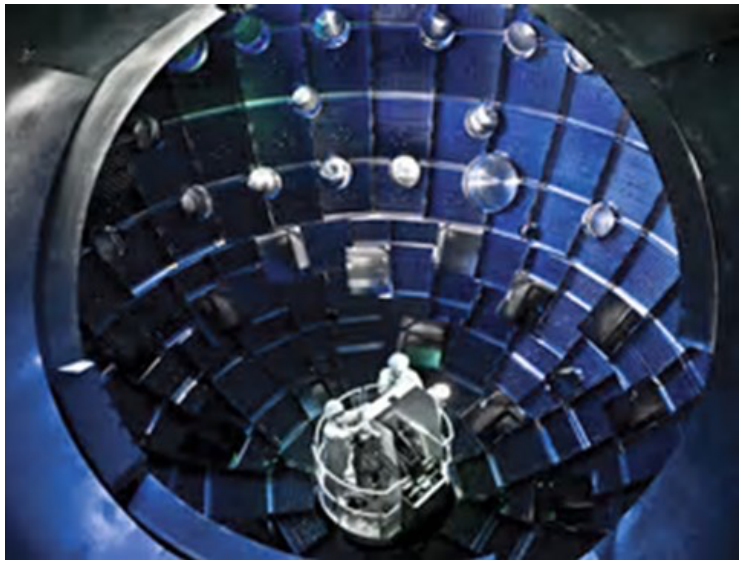
Our Technology World is Impressive

Amazon.com Distribution Centers

Our Technology World is Impressive

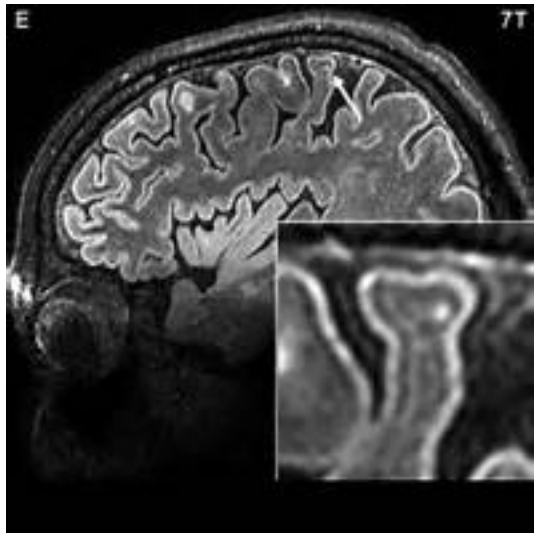
Facebook data center Lulea, Sweden (Artic Circle)~ 300 million US$
Facebook process 350 million photographs, 4.5 billion “likes,”
and 10 billion messages a day. Refrigeration power ratio 3--> 1.04

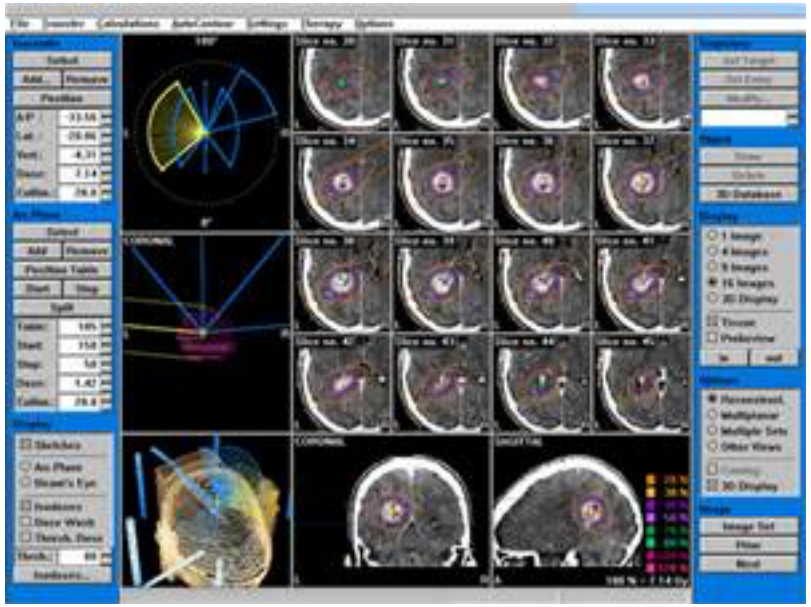
Our Technology World is Impressive



US National Ignition Facility, Lawrence Livermore Labs
total energy is amplified 10^{18} from 1 billionth of a joule to 4 million joules
—all in a few millionths of a second

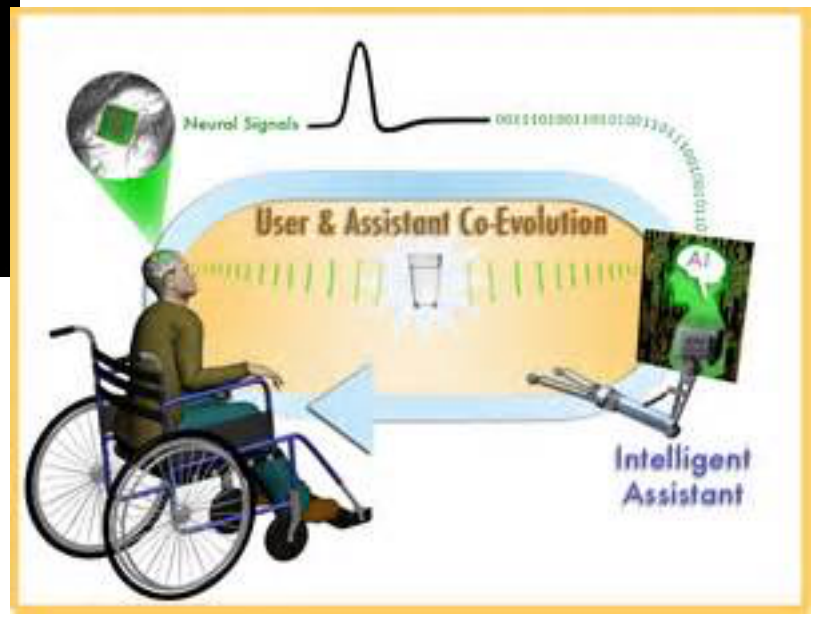
Our Technology World is Impressive

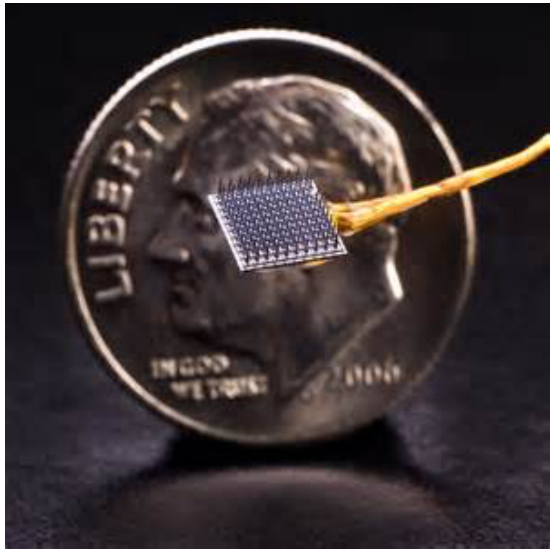




Robotic Surgery and the new medical Diagnostics

Our Technology World is Impressive





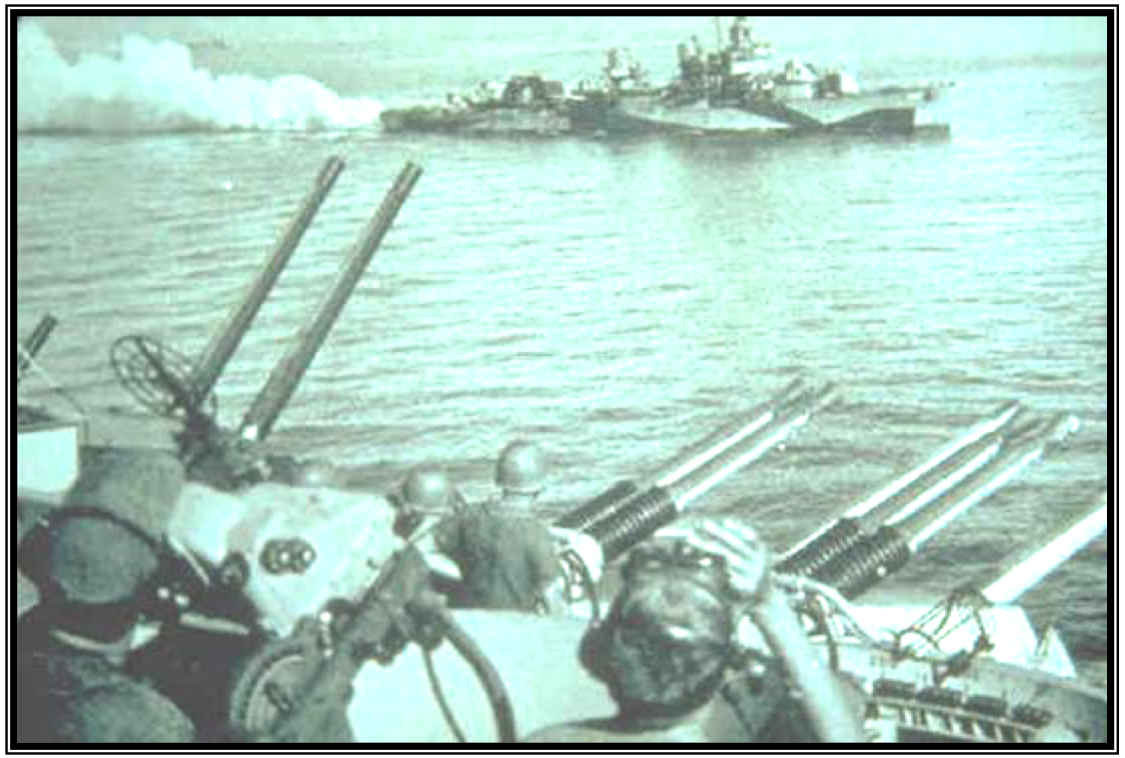
Brain Computer Interfaces to help the disabled

Trend to Brain Centric Technologies

- Our brain is an internal representation system.
- Life in society is difficult because requires brain to brain communication
- Throughout history we invented ways of copying with this problem: culture, language, books, and now the web repositories.
- AI is going to bridge brains and create an external intelligent society to facilitate access and diffusion of information amongst humans.
- This will drastically accelerate changes in the way we live and organize society.

Master Slave Tools: The Dawn of a New Era

The anti-aircraft fire control 1941

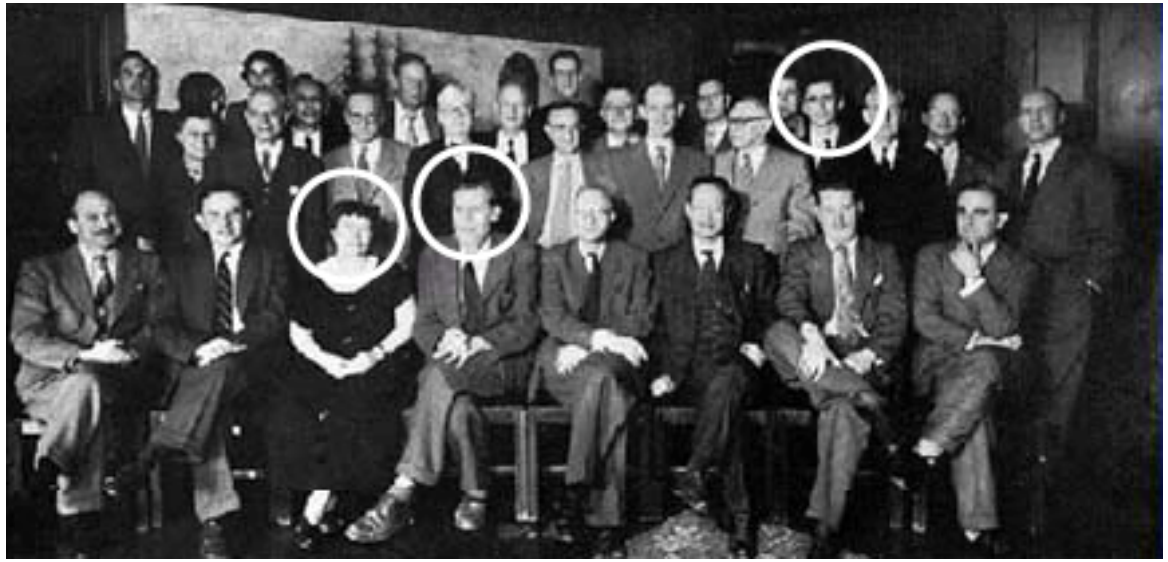


Machine could automatically track a target and predict its position in the future (Norbert Wiener/Julian Bigelow)

After The War…
Josiah Macy Foundation Meetings (1946-53)

Participants:

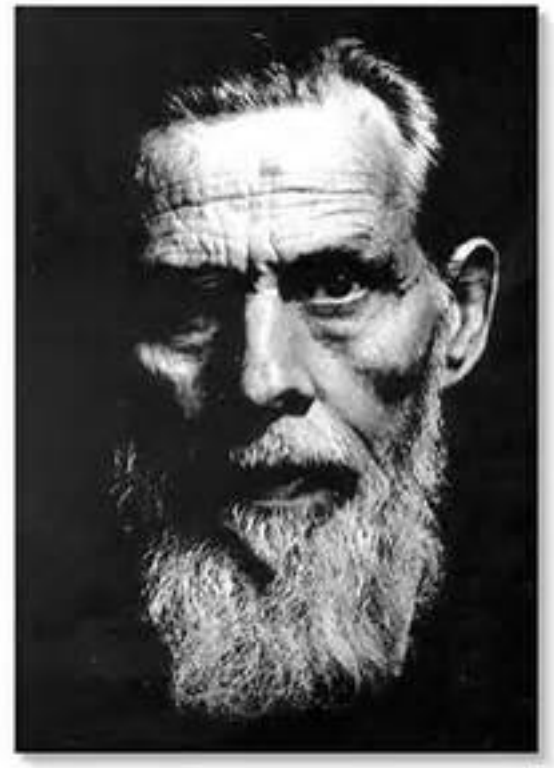
McCulloch (Chair)
Wiener
Von Neumann
Pitts
Shannon
Margaret Mead



- Ideas from:
  - Living organisms
  - Control and communication devices
  - Human society
- Focus on:
  - Information, feedback, control= cybernetics

Warren McCulloch Neurophysiologist

(1943)



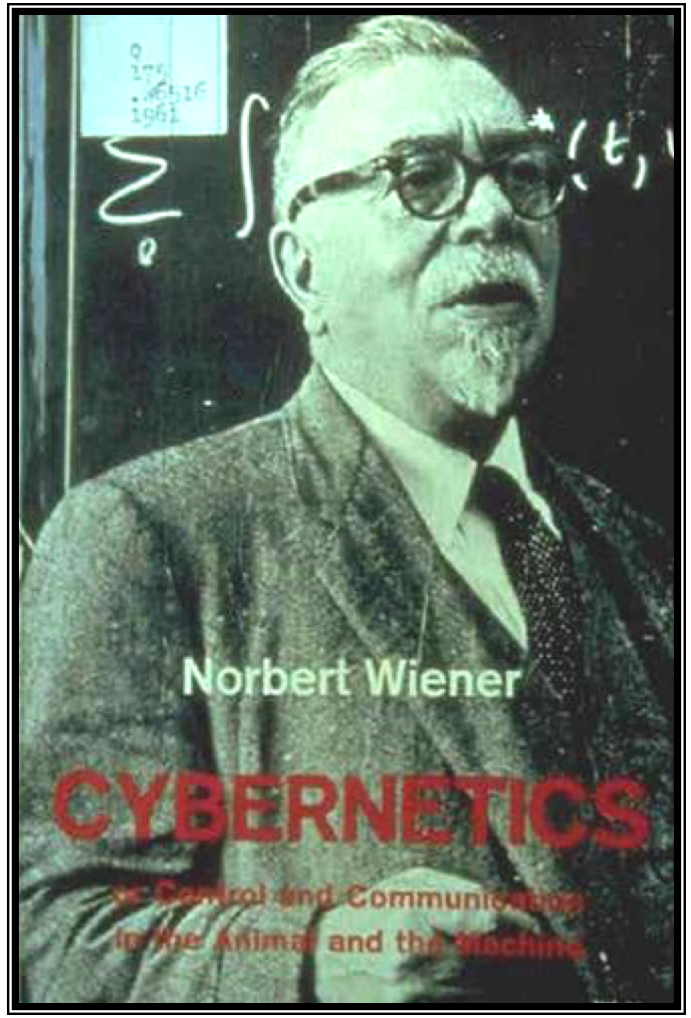
- Implementation of Turing Machines in brain like components (McCulloch-Pitts units)

- Started the neural network revolution
- Cybernetics
- Book: Embodiments of Mind , MIT Press 1965

McCulloch W., Pitts W.,"A Logical Calculus of the Ideas Immanent in Nervous Activity" Bulletin of Mathematical Biophysics Vol 5, pp 115–133, 1943

Norbert Wiener –Mathematician

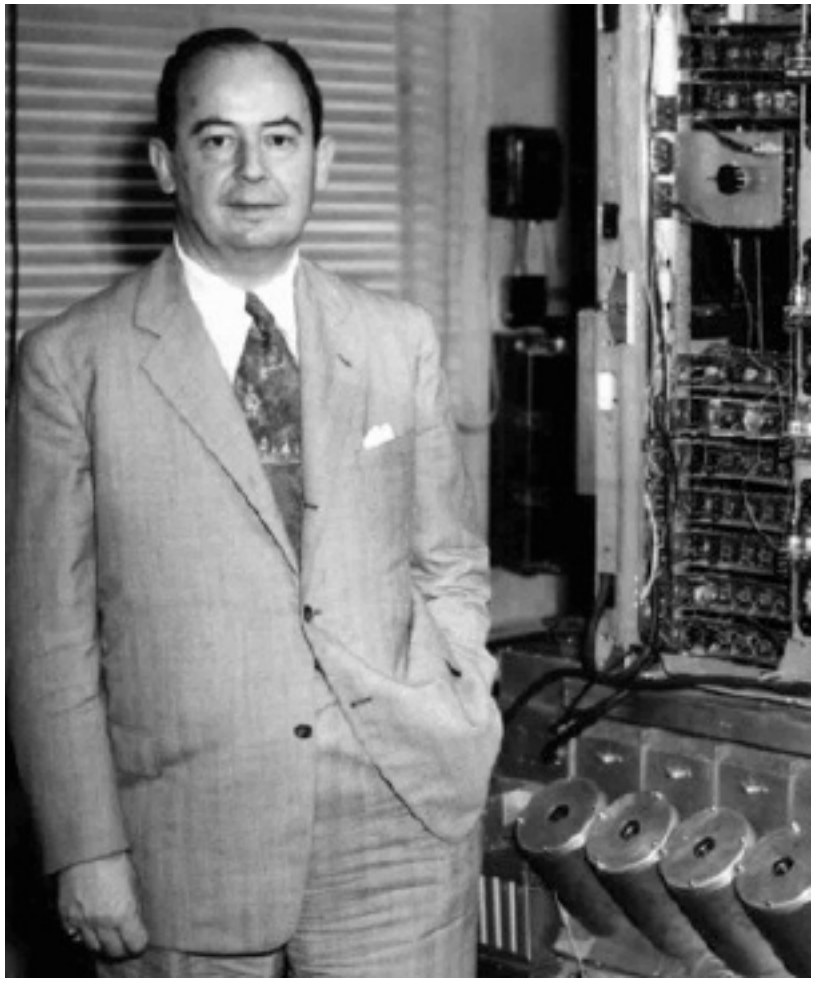
Cybernetics (1948)



Wiener believed that there were mathematical principles that explained human behavior

- Systems that have goals
  - Using circular “feedback” loops of
    - sensing→comparing with goals →action → sensing …
  - Understand human interaction to design computer interaction

John von Neumann–Mathematician Inventor of Digital Computer (1952)



- Implementation of Turing machine in digital hardware
- Studies on complex automata made from unreliable parts

Quote from 1951

A new, essential logic, theory is called for in order to understand the high complexity automata and in particular the central nervous system. It may be however that in this process, logic will have to undergo a pseudomorphosis to neurology to a much greater extent than the reverse.

Claude Shannon
Electrical Eng.

(1948)

- Invented Information theory, so he is the father of the information age!
- New branch of statistics
- Optimal design of communication systems to achieve error bounds in arbitrary channels

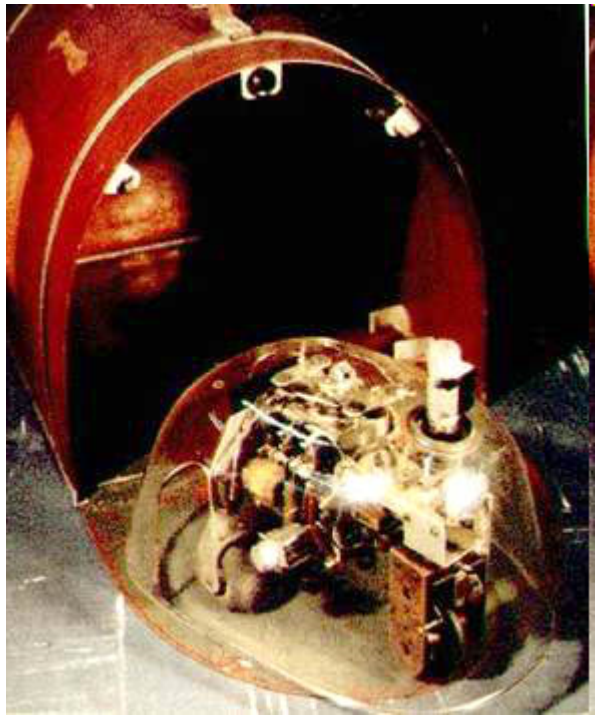
Claude E. Shannon: A Mathematical Theory of Communication, Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, 1948.

W.G. Walter
Neurophysiologist and Roboticist
(1950)

- Neurologist who did computer analysis of EEG with Fourier analyzers

- Bristol Turtoise, autonomous robots with sensors that would explore the environment and seek light to recharge their solar cell batteries

- Seven postulates for autonomous behavior



Walter W.G. The living brain , 1953

Jo C. R. Licklider
Psychologist (1960)

- Internet pioneer

- Postulated “man-computer symbiosis”
- Tightly coupling human brains and computing machines to revolutionize information handling
- Concept of a “dialogue” and “partnership”

Licklider, J. C. R., Man-computer symbiosis, *IRE Transactions on Human Factors in Electronics*, *HFE-1*, 1960, 4-11.

Frank Rosenblatt

Psychologist (1956)

- Invented the Perceptron,
- The first learning machine that could generalize

- Proof of convergence

Rosenblatt F., Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms", Spartan Books 1962.

Donald O. Hebb
Psychologist (1949)

- Propose the Hebbian learning rule, for synaptic plasticity
- Still the only widely accepted learning rule in computational neuroscience

Bernard Widrow
Electrical Engineer

- Inventor of the LMS (least mean square) algorithm and the father of adaptive signal processing.
- Widely applied in modems (channel equalization)
- Work on ADALINE

Jeoffrey Hinton
Computer Scientist

- Co-inventor of the backpropagation algorithm.
- Boltzmann machines
- Very influential (postdocs, ph.d.) and the godfather of deep learning

Paul Werbos
Econometrics

- Inventor of the Backpropagation through time algorithm to adapt recurrent networks,
- Very influential figure at NSF to fund neural network research

Stephen Grossberg
Computational Neuroscience

- Early contributor in neurodynamics,
- Very influential – EiC of Neural Networks, founder of International Neural Network Society
- Propose the Adaptive Resonance Theory for unsupervised learning

David Rumelhart
Psychologist

- One of the proponents of the backpropagation algorithm,
- Very influential proponent of the parallel distribution processing in cognitive science

Teuvo Kohonen
Electrical Engineer

- Early contributor to associative memories
- Developed the learning vector quantization algorithm
- Proponent of the Self Organizing Map

Kunihiko Fukushima
Computer Scientist

- proposed the NeoCognitron, 1980
- Very influential because it was a deep architecture (precursor to CNNs) that was trained in a supervised/unsupervised manner

Michael I. Jordan
Psychology & Statistician

- Early contributor, in artificial recurrent networks
- Very influential computer scientist

Vladimir Vapnik
Statistician

- Very influential work on a new form of statistical learning called Structural Risk Minimization
- Inventor of the support vector machine (SVM)

Yiann Lecun
Computer Scientist

- Earlier contributor to neural networks
- A major proponent of Deep Learning
- Very influential (postdocs, ph.d.)
- Developed the Convolution Neural Network

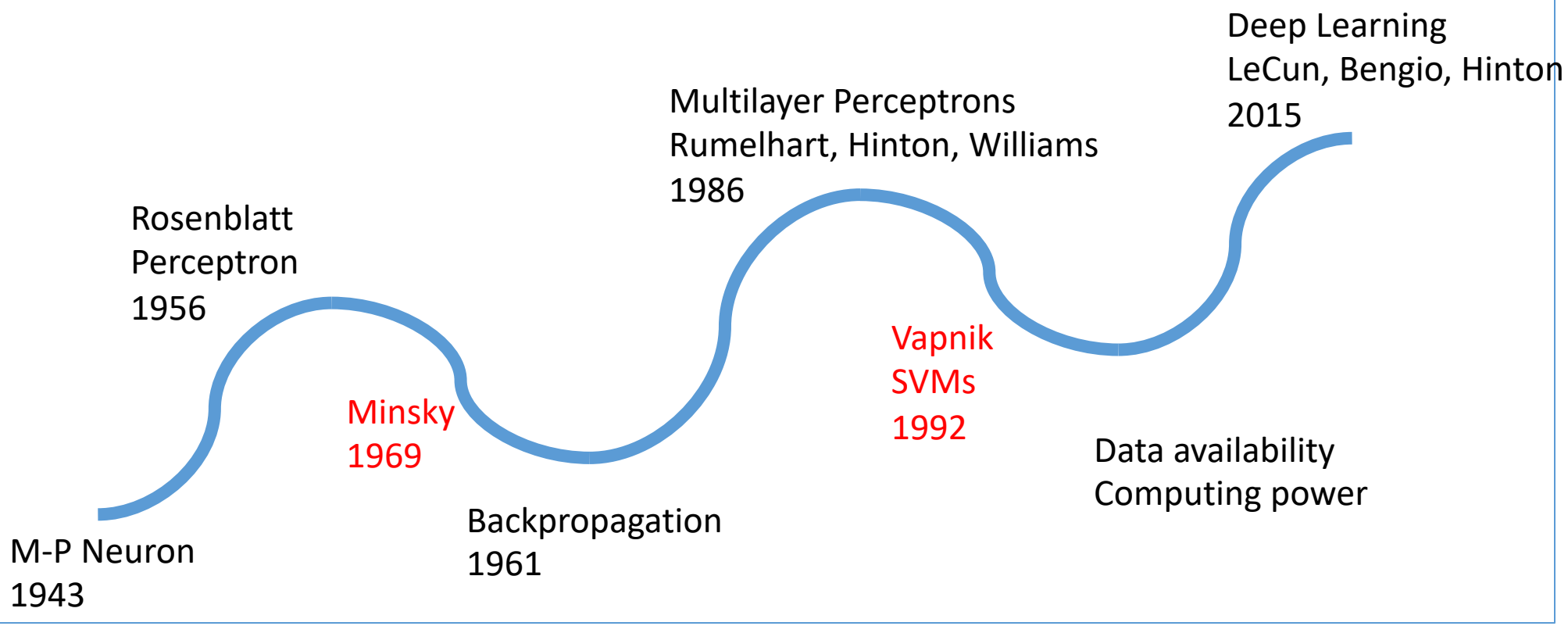
- Chief AI Scientist at Facebook

Jurgen Schmidhuber
Computer Scientist

- Early contributor in neural networks
- Very influential (postdocs, ph.d.)
- Inventor of the LSTM (long short-term memory)

Cycles in Neural Network Research

- How to manage expectations?



Explaining the last Upswing

- The latest upswing was not caused by new theories: it was due to the availability of large data sets, fast computers and the interest of the AI industry.
- Speech recognition and object recognition, both perception tasks, were unmet needs for over 50 years.
- With large data sets and huge computer clusters Microsoft, Google, Baidu and Facebook using deep learning were able to crack these two big fields and will reap the beneficit.

Looking at the Future

- The methodologies for speech and face recognition are well established, but still use unrealistic methods to solve the problem (supervised learning).
- A general way to perceive the world using machine learning for autonomous robotics and commercial use is still lacking. I am sure it will come, but is not here today.
- And for those of us that understand the cognitive brain, perception is the “easy piece”, so there is a long way to go.
- So hype, once again, may kill the momentum.

Looking at the Future

- In my research I am interested in **man-machine symbiosis**
- The opposite of the popular press, which fears a point where machines can outperform our intelligence.
- The goal is to emulate the human perception-action-cycle in machines and let them use our physiological data and intentions to help us solve the problems we face everyday.
- We already have most of the pieces, but the architecture for software integration is lacking, and machine learning will be part of it.

Conclusions I

- We are at a cross road in scientific methods: deductive reasoning gave us our current understanding of the world.
- Machine learning is an inductive process, so there are few guarantees of performance in new data (look at the prognostic in the recent US election!).
- Contrast this to engineering design (also an inductive process): understand from first principles what works, how systems fail, put pieces together, and create an envelope of performance.

Wigner E., 1960 “Unreasonable Effectiveness of Mathematics in the Natural Sciences”, *Comm. on Pure and Applied Math*. **13**: 1–14
Hamming R., 1980 “Unreasonable Effectiveness of Mathematics”, *The American Mathematical Monthly*. **87** (2): 81–90
Pereira F., Norvig P., Halevy A., 2009 "The Unreasonable Effectiveness of Data", *IEEE Intelligent Systems*, vol. 24, no. , pp. 8-12

Conclusions II

- Complexification of society and processes is here to stay
- What matters is not the big data we collect but the **information that the data contains**.
- Information is the key commodity for decision making because it has a practicality that data lacks.
- Information Theory will have a leading role in the future because it quantifies uncertainty, and together with machine learning can lead to optimization of information content.
- When we change the information flow in a complex system, we change the system, so this is also the best tool for management.

Principe J., Information Theoretic Learning, Springer Inf. Science and Statistics, 2010.