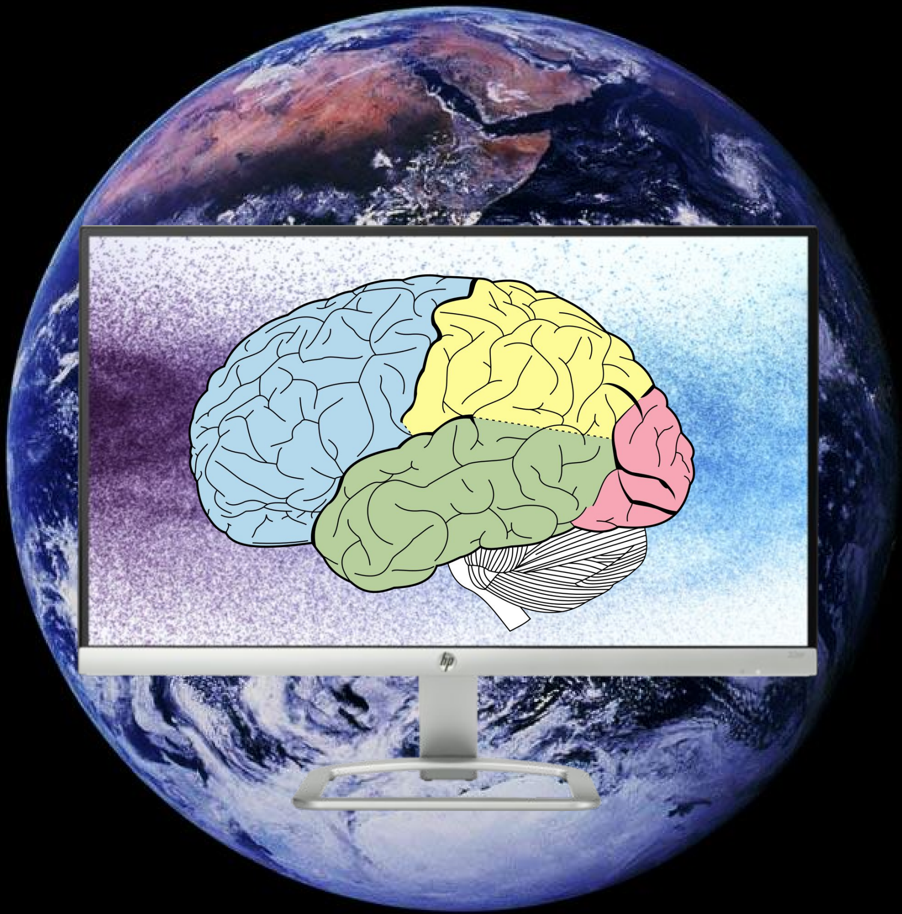
KEXP

KNOWLEDGE EXPLORATION SYSTEMS

Danbury AI

Artificial Intelligence & the Dawn of **Deep Learning**



Andrew Ribeiro 1/2/2018

Talk Outline

AI and the Dawn of Deep Learning

- About Me and My AI Journey
- Life and Intelligence
- AI History and Perspectives
- Neural Networks
- Applications of Neural Networks and Deep Learning
- Challenges in Neural Network Research

Mathematical Section *(optional)*

- Forward propagation of a neuron.

Conclusion

- Getting into AI
- Questions

AI and the Dawn of Deep Learning

About Me and My AI Journey

Where I've been

KEXP

About Me

- Co-Founder of Knowledge-Exploration Systems.
- Co-Organizer of Danbury AI **April 2016**
- Studied Computer Science at WCSU
- Main interests
  - Machine Learning
  - Natural Language Processing
  - Adaptive Learning Systems
  - Collaborative Technology
  - Theoretical Computer Science
  - Philosophy and Psychology
  - Mathematics (Discrete, Linear Alg, Foundations)

My Motives for Studying AI

- Philosophy, psychology, mathematics, and computer science are my core interests. There is no field other than artificial intelligence that brings these subjects together under such a cohesive and exciting banner.
- I believe the program of analytic philosophy ends up in AI.
- AI is a difficult problem that requires lots of tools and tricks that extend into almost all areas of mathematics. Being focused on applications to AI helps me stay focused on particular areas of mathematics that have utility in studying the problem. Instead of studying pure mathematics in a void, I always ask the question “How can this be applied to AI?” And this helps me form a more tangible relationship with pure mathematics.

My Motives for Working in AI

- It’s an exciting field that is growing rapidly with constant new advancements and research that has the potential to change many industries.
- People are turning to AI to solve some of the most interesting and challenging uses of computers, and I want to see how far computation can go while helping press the boundaries of computation.
- People working in AI are generally highly adept at mathematics and CS and I like that environment as opposed to standard dev houses.
- There are some problems I’d like to work on that require some form of modeling intelligence. (Mainly NLP and different knowledge based tasks like adaptive learning systems.)

How can we use AI to....

Improve the way we learn?

Accelerate scientific research?

Connect people by facilitating peace and teamwork?

Better utilize computation in daily life?

Improve our creativity and productivity?

Answer fundamental questions about our nature?

ABOUT DANBURY AI - *Mastering AI Together*

Danbury AI is a public AI meetup group hosted by the Danbury Hackerspace that aims to stimulate discussion in the vast field of artificial intelligence and bring together locals that foster a passion for the field. We hold our in person meetings on the first Tuesday of every month at which we host a variety of presenters, discussions, and workshops. We maintain a lively web presence via slack so the conversation never stops.

"If you want to go fast, go alone; If you want to go far, go together."

- Heavily inspired by initiatives like Open AI.
- We hold monthly **in-person meetings** at the Danbury Hackerspace.
- We aim to conduct **study groups** and other initiatives of group learning. We have a wide variety of members from all reaches of industry.
- We have a group chat room and **forum on slack for sharing resources and coordination.**
- Main website: https://www.meetup.com/DanburyAI/

DANBURY AI: Organizers - *Mastering AI Together*

Andrew Ribeiro

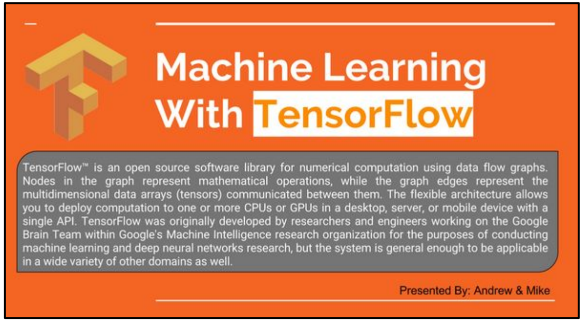
Western Connecticut State University, ND. Computer Science

Michael Rogowski

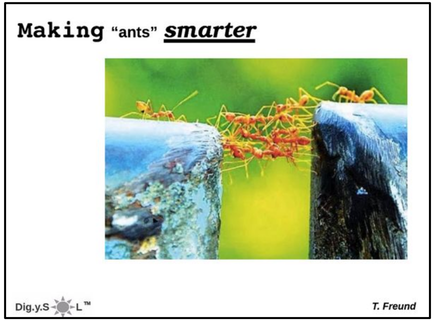
Western Connecticut State University, BA. Computer Science

Lambert Wixson

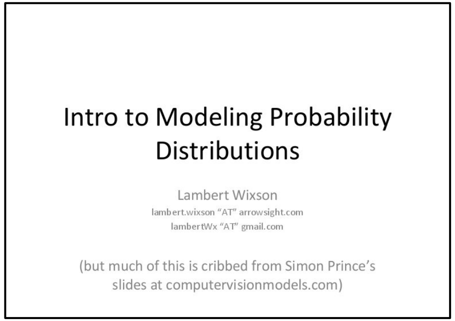
University of Rochester, PhD. Computer Science



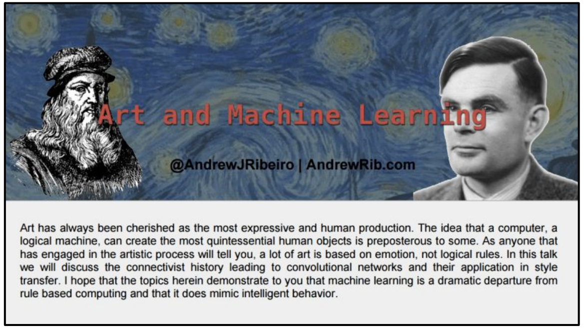
Slides | Repo | Workshop



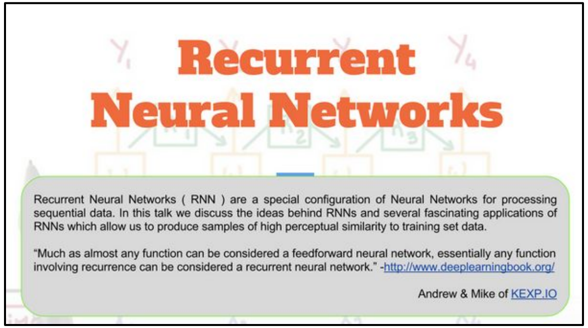
Slides



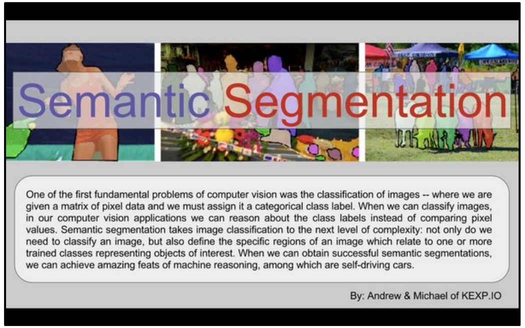
Slides



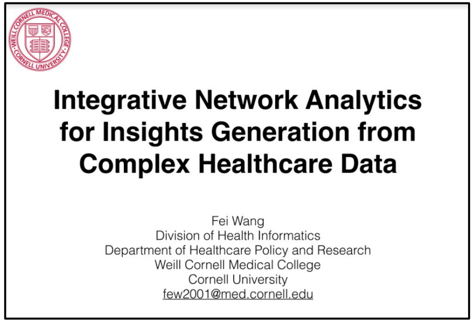
Slides



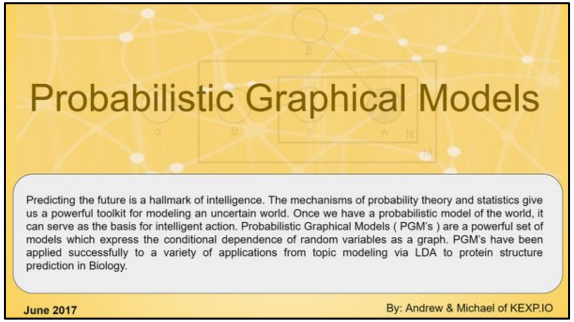
Slides



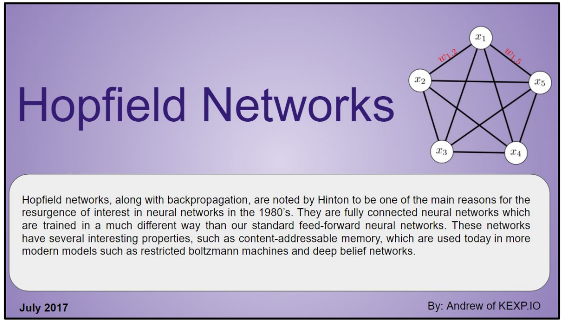
Slides



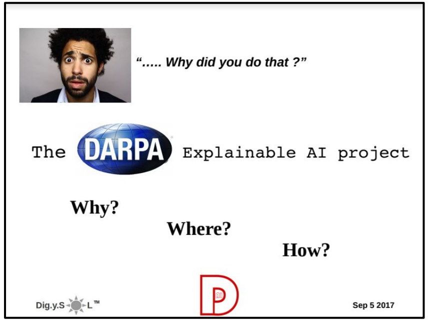
Slides



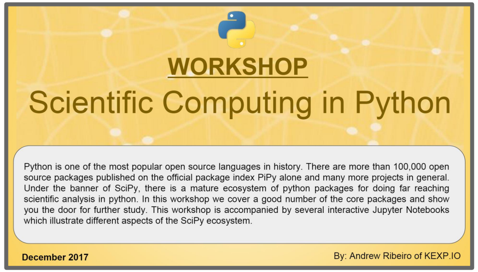
Slides



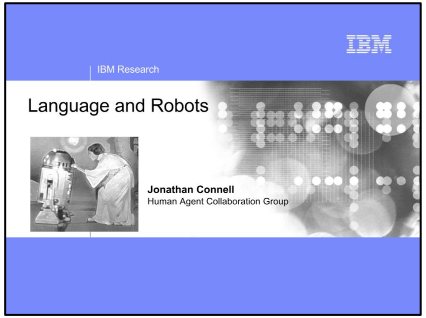
Slides



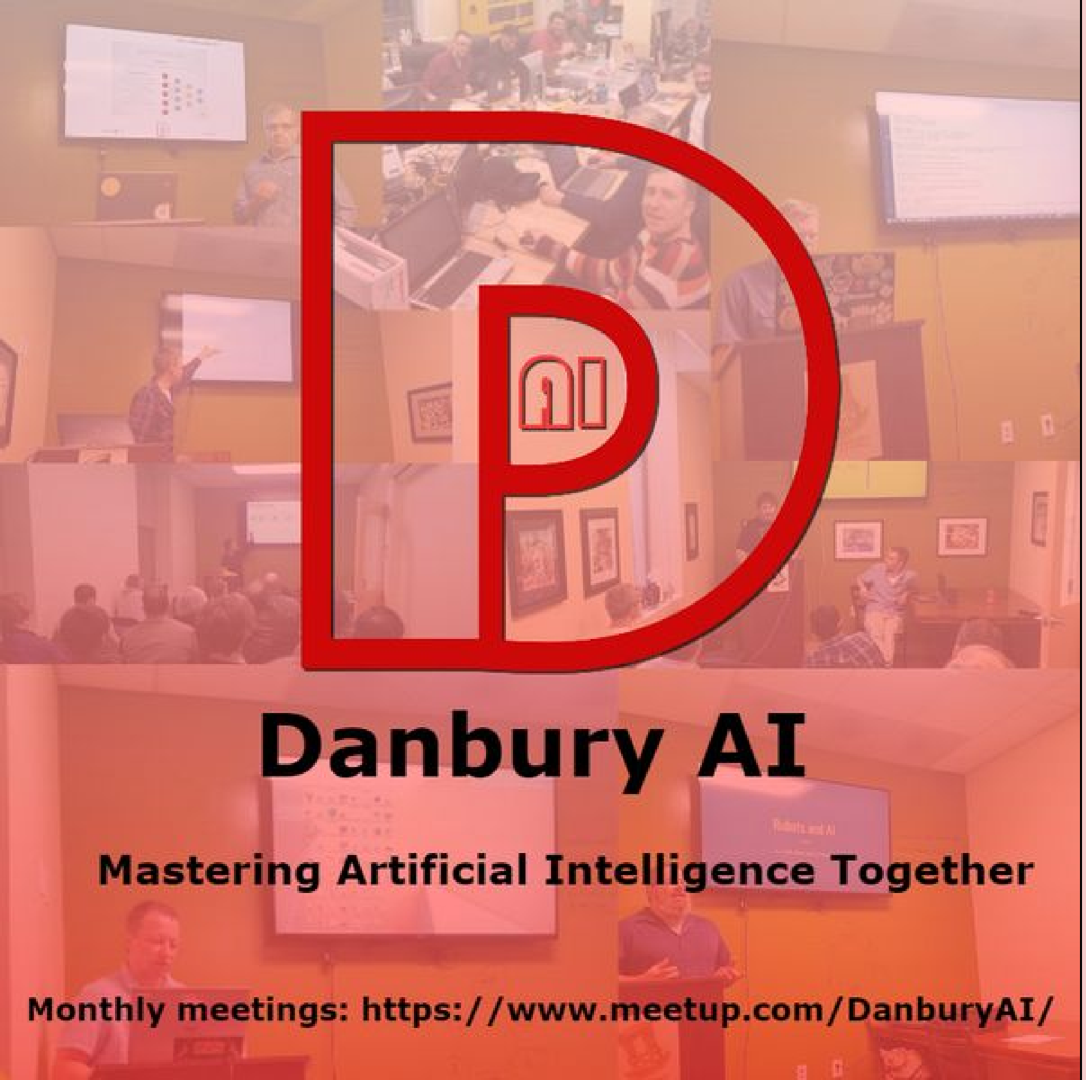
Slides



Repo | Slides



Slides | Video



Interested in AI?
A master of AI?
A student of AI?

Join Us

DANBURY
HACKERSPACE

We're building a community of makers, artists, craftspeople, hackers and entrepreneurs, and we are looking for people to join our group. Our home is in the old Union Savings Bank building connected to the Danbury Library at **158 Main Street**, and we have a program space, coworking room, mockup studio, and community area with a 21' custom-built maple table. Our tools include 3D printers, a CNC router, wood and metal-working tools, and more. Membership is $50 per month but we're always willing to give tours and mentor local entrepreneurs.

Danbury Tech Groups

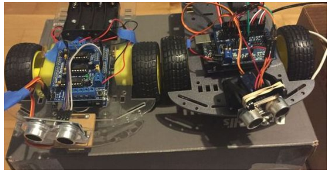
Danbury Artificial Intelligence

Danbury Hackerspace

danbury.io

Danbury Area Computer Society

Danbury Engineers of Robotic Platforms



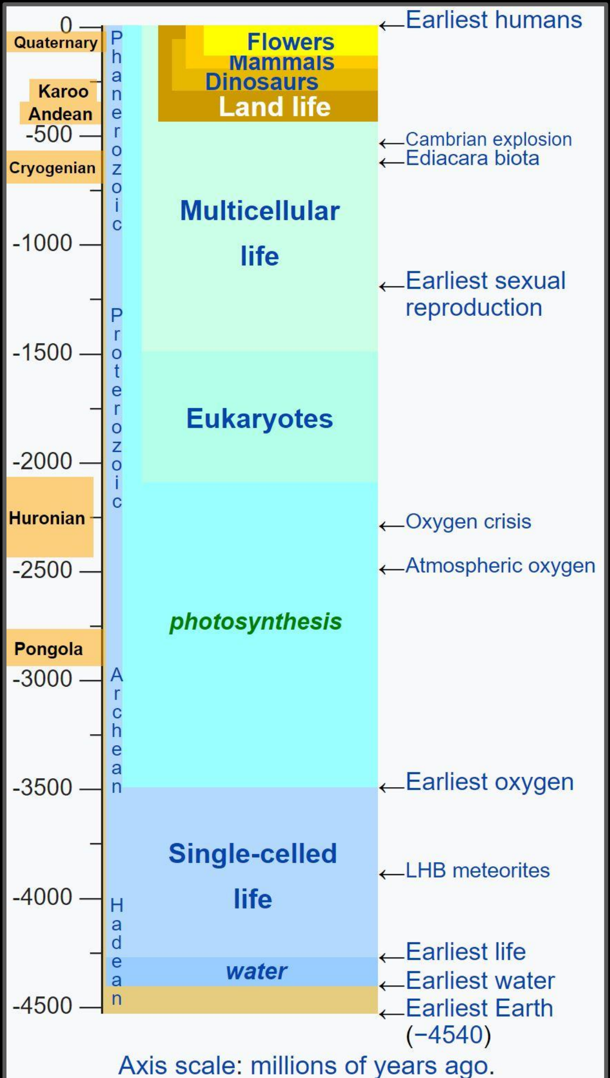
KEXP
KNOWLEDGE EXPLORATION SYSTEMS

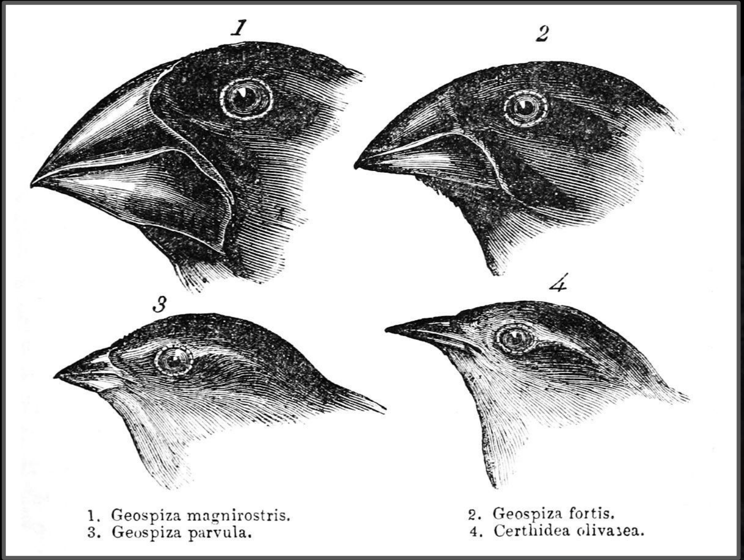
- Founded in **April 2016** by Mike and Andrew
- How do we create a company around modern AI research and development?
    - Pythonic Machine Learning and Scientific Computing
    - Cloud Computing
    - Full Stack Development.
- Currently seeking more B2B contracts.

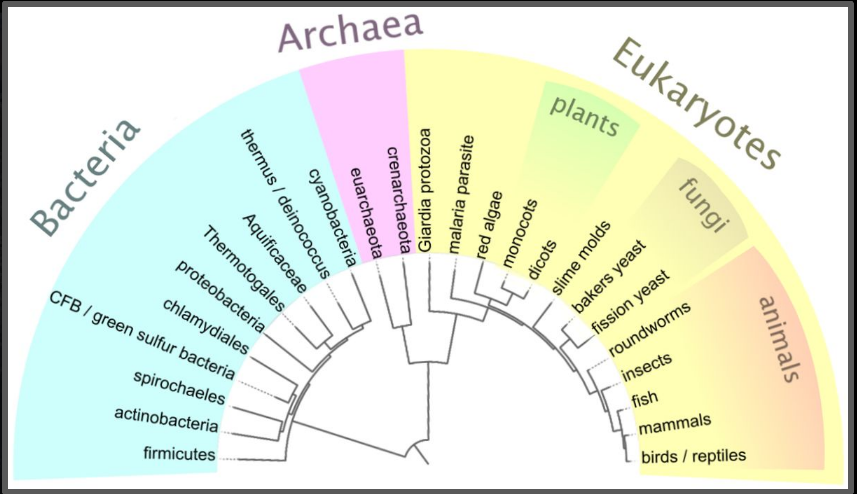
Life and Intelligence

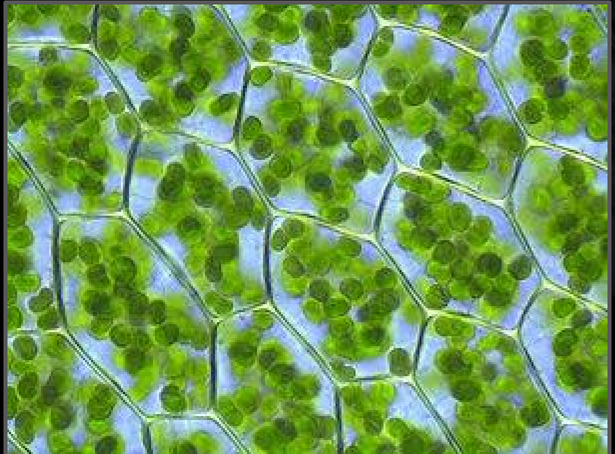
Life

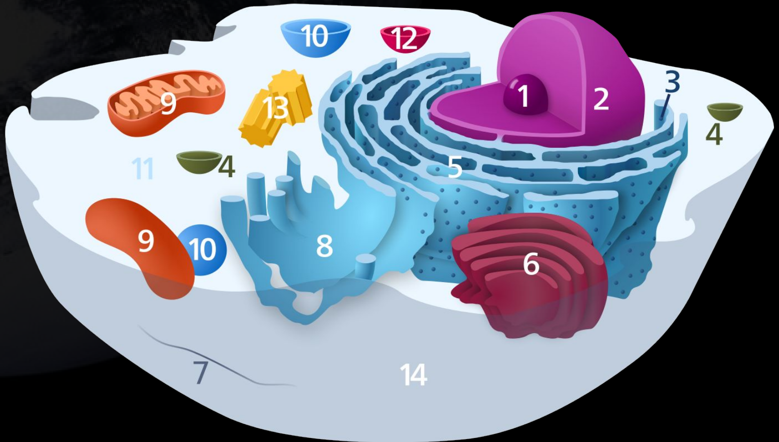
- Life, through evolution, is naturally selected to make efficient use of resources in its environment.
- The physical world itself is biased toward parsimony. From the first principles of physics, see the principle of least action, to chemistry, to life.
- As life co-evolves, organisms use other organisms to increase their resource utility. This is the basis of multicellular life, ecology, society, and the brain.
- The universe tends toward entropy, while life tends towards complexity and greater information. Life itself is self-replicating information.
- The diversity of life comes from the diversity in resources and the way each lifeform evolves to take advantage of them.




1
2
3
4
1. Geospiza magnirostris.
2. Geospiza fortis.
3. Geospiza parvula.
4. Certhidea olivasea.


Bacteria
Archaea
Eukaryotes
plants
fungi
animals
CFB / green sulfur bacteria
spirochaeles
actinobacteria
firmicutes
chlamydiales
proteobacteria
Thermotogales
Aquificaceae
thermus / deinococcus
cyanobacteria
euarchaeota
crenarchaeota
Giardia protozoa
malaria parasite
red algae
monocots
dicots
slime molds
bakers yeast
fission yeast
roundworms
insects
fish
mammals
birds / reptiles




1
2
3
4
4
5
6
7
8
9
9
10
10
11
12
13
14

On Intelligence in General

- Why did intelligence evolve in living things?
  - Prior to the cambrian explosion, most organisms were composed of single cells. As organisms became increasingly multicellular, there became a pressing need for coordination and specialization. In animals, this coordination system evolved to what we know as the nervous system. The brain is the key component in the nervous system.
  - Sponges are the only multicellular animals without a nervous system.
- Why did a mind evolve?
  - It is clear from experience that not all animals exhibit what we call a mind. Small animals, which often have simpler nervous systems, sometimes don’t get to the point where they have the hardware to support mind.exe.
  - The mind seems to have developed as a simulation in which the consequences of actions can be evaluated. It costs nothing to perform a poor action in the mind, but it can be quite costly if a poor action is made in the real world. Thus the mind is an evolutionary advantage.

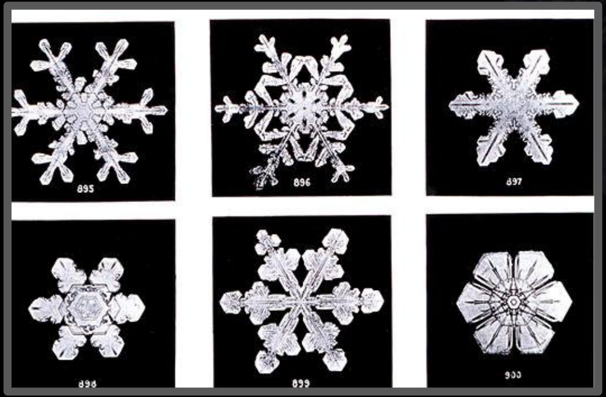
Constructionism

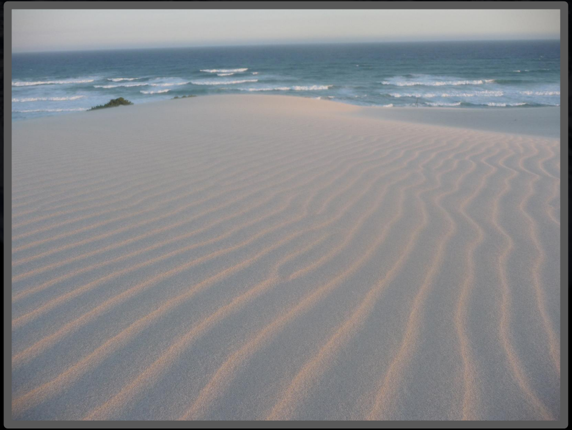
- "The word constructionism is a mnemonic for two aspects of the theory of science education underlying this project. From constructivist theories of psychology we take a view of learning as a reconstruction rather than as a transmission of knowledge. Then we extend the idea of manipulative materials to the idea that learning is most effective when part of an activity the learner experiences as constructing a meaningful product." - Papert 1987
- Constructionism has parallels in the philosophy of mathematics.
- Homotopy Type Theory (HoTT) is a type of intensional type theory focused on constructing mathematical objects. This field represents some of the finest modern mathematics done at the Institute for Advanced Study.

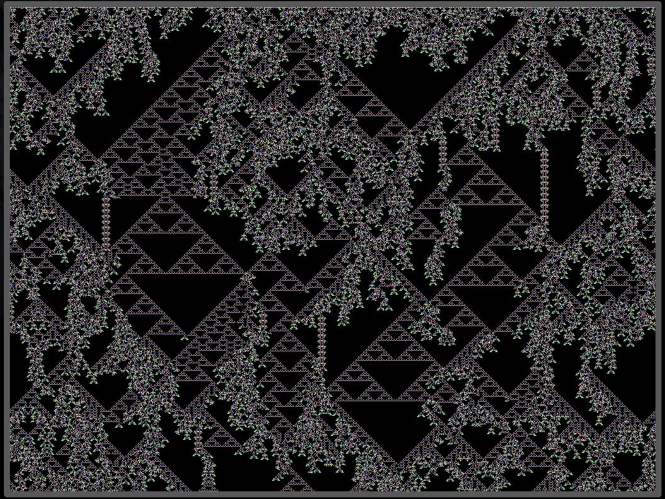
Emergence and Complex Systems

- “A complex system is a system composed of many components which may interact with each other. In many cases it is useful to represent such a system as a network where the nodes represent the components and the links their interactions.”
- “Emergence is a phenomenon whereby larger entities arise through interactions among smaller or simpler entities such that the larger entities exhibit properties the smaller/simpler entities do not exhibit.”
- One approach to modeling complex systems is to find simple rules that govern simple components of a complex system; which when combined together produce complex emergent characteristics.


Fourmilière
9
Cocons de fourmis
Cocons de fourmis


895
896
897
898
899
900





AI Perspectives and History

AI Perspectives and History - Overview

- Defining AI
- Problems and Scope of AI
- Black Box Intelligence
- The Turing Test
- Chinese Room Argument
- AI by definition
- Mathematics
- Applied Mathematics
- Mathematics <--> Applied Math
- AI as Logic, Language, and Symbolic Reasoning

- AI as Control/Cybernetics and Mechanics
- Early Homeostatic Devices
- AI as Search
- AI as Optimization
- Machine Learning
- Why Today's AI Starts with Data

Defining AI

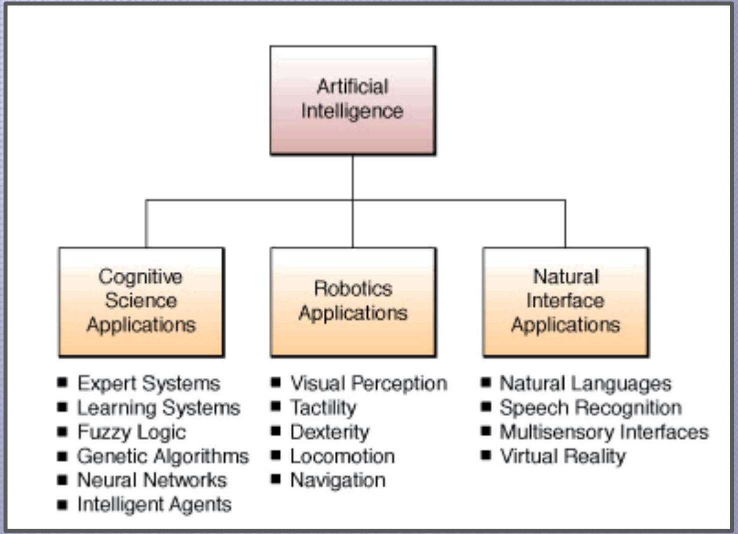
“As soon as it works, no one calls it AI any more.” - John McCarthy

- AI: “It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.” -John McCarthy
- **Strong AI**: “The appropriately programmed computer with the right inputs and outputs would thereby have a mind in exactly the same sense human beings have minds.” - Searle.
- **Weak A**I: A weak AI simply mimics a human and does not have internal processing isomorphic to human minds. (See chinese room argument.)
- A strong AI is said to be an **Artificial general intelligence (AGI)** because it can perform any intellectual task that a human being can
- A big problem in defining AI is that it makes use of the concept of intelligence, which still is not rigorously defined. Only when we have a computational model of intelligence, AI, do we truly have a rigorous definition of intelligence.

My definition: AI is the investigation of intelligence with mathematics.

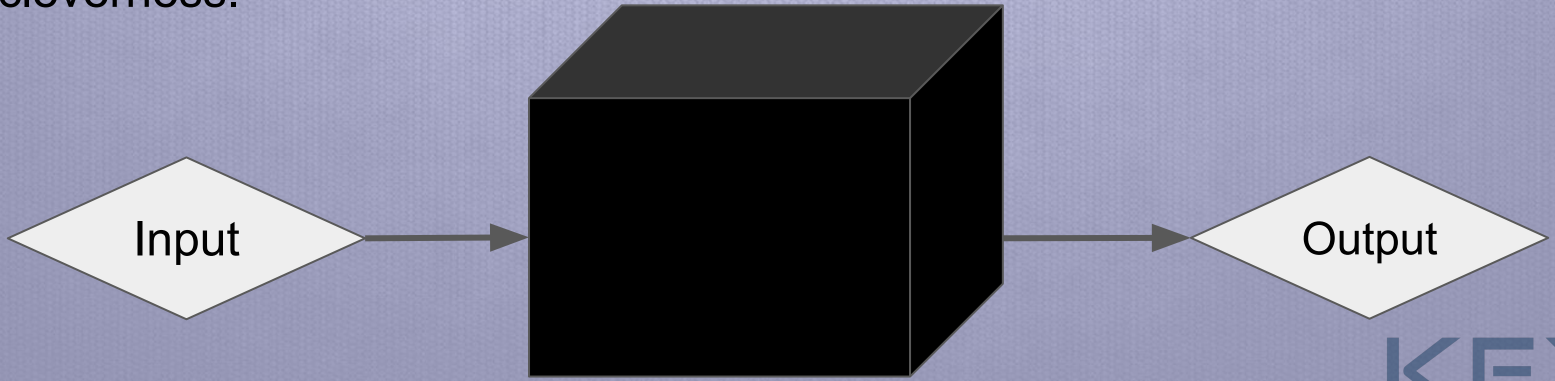
Problems and Scope of AI

- Reasoning and Problem Solving.
- Knowledge Representation and Modeling.
- Learning
- Planning
- Perception
- Prediction
- Motion and Manipulation
- Creativity



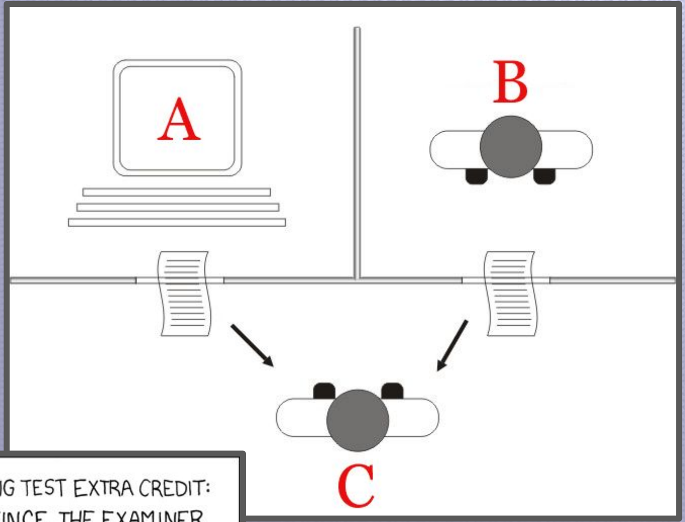
Black Box Intelligence

- If we ignore all the nuances of modeling intelligence, the simplest model of an entity is a black box with an input and output.
- If what's in the box changes over time in respect to input, meaning it produces different outputs in respect to inputs, and the **error** of that output can be minimized over time, we may say that the entity can learn.
- We may deem this box intelligent if the output is sufficiently clever in respect to the input. The Turing Test is one such test of sufficient black box cleverness.

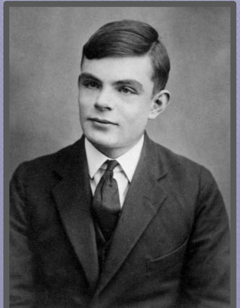


The Turing Test - The Imitation Game

Turing posed that the types of machines that would pass the imitation game would be machines that learn.



- Introduced by Alan Turing among several other ideas in his 1950 paper "Computing Machinery and Intelligence."
- The identity of A or B is unknown to the interrogator C. C must determine between A or B which is a computer and which is a human.
- The Turing test is considered a black-box test of intelligence, in that we cannot peer into the construction of A nor B. We must judge it strictly on output in respect to inputs.



John Searle

Chinese Room Argument - Competence without comprehension

Is the entity inside the room really intelligent? Or are it simply following a set of rules prescribed by an outside intelligence? This can imply that the Turing test is not a sufficient test for strong AI, only weak AI.

KEXP

AI by Definition

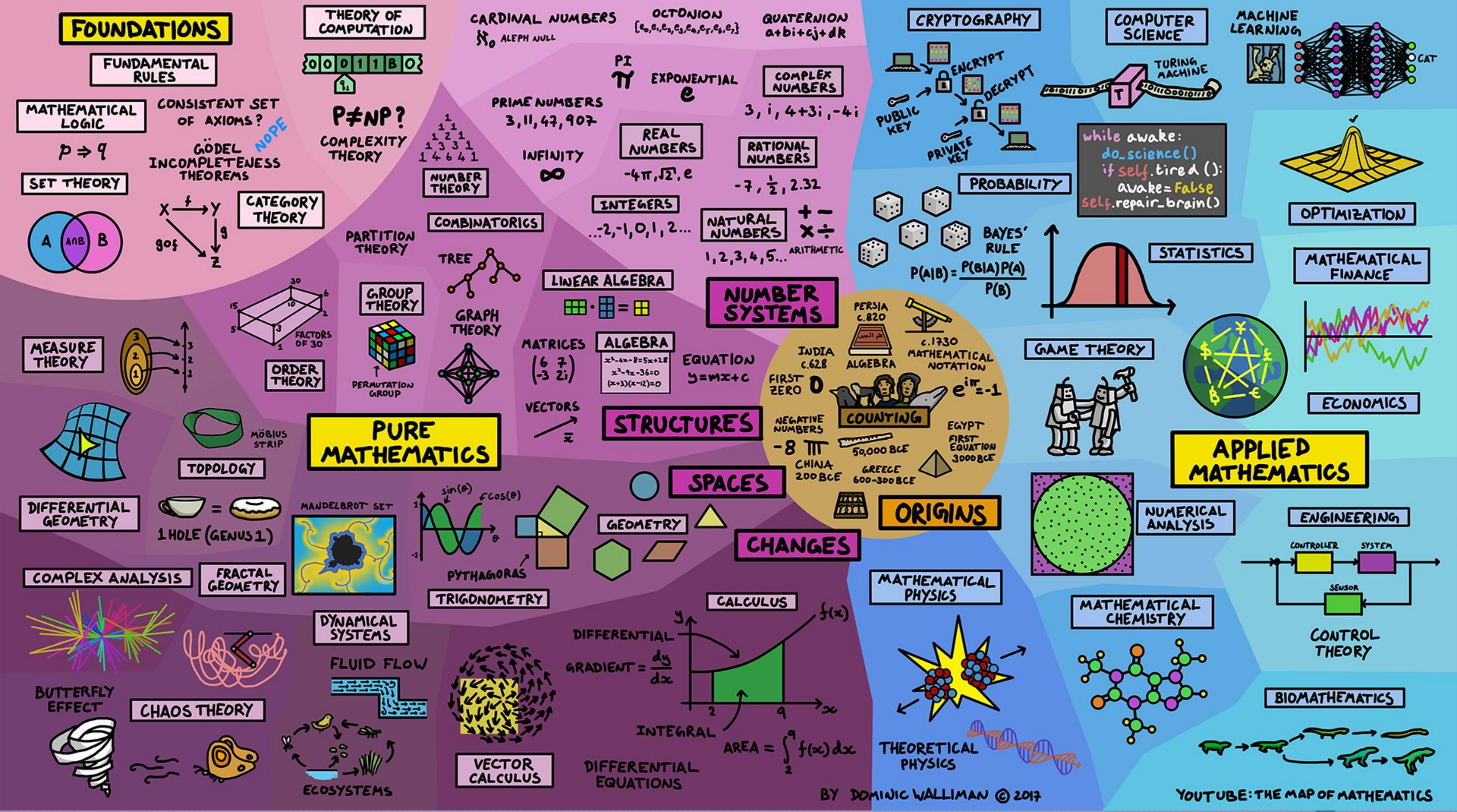
- Instead of relying on black box tests of intelligence, AI by definition argues that an entity is intelligent with respect to another entity by the degree to which they are mechanically similar. This perspective admits that the mind is mechanical.
- Machines that do not learn lack a key mechanism that human beings and other animals poses. Therefore only machines that can learn are candidates for strong AI.
- Neural Networks not only can learn, but they learn in way that is biologically inspired. Neural Networks are our closest candidate for AI by definition, but we still haven’t figured out how to solve general AI with them.

Mathematics

The **rigorous** study of **structure** divorced from details by way of **abstractions**.

- **Rigorous**: Results can be proven. (Induction, axiomatic deduction, etc)
- **Structure**: The matter of the subject that forms its identity apart from other things. “ The arrangement of and relations between the parts or elements of something complex.”
- **Abstractions**: Representations of structure that omit details seen to be insignificant to understanding its nature wrt to our modeling objectives.

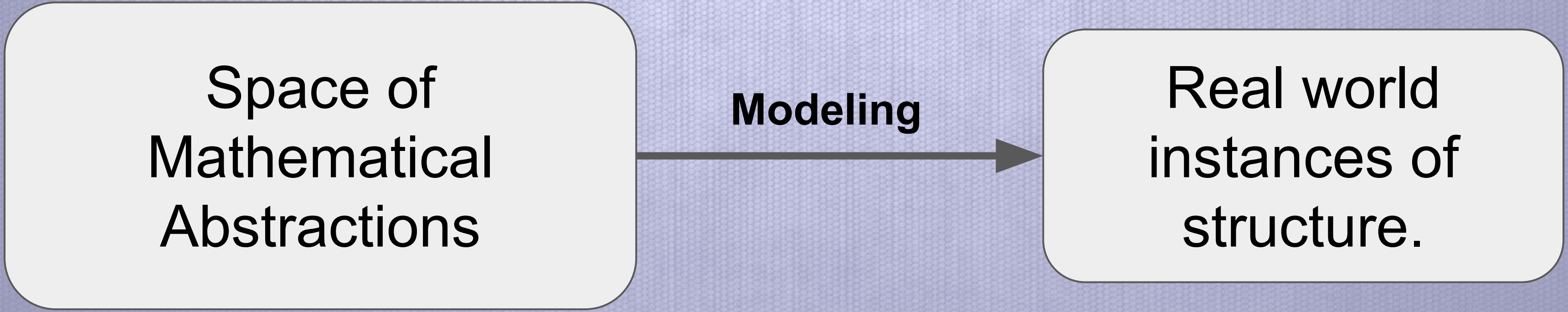
Numbers offer a convenient and natural way of quantifying the nature of structures in an abstract manner. Please don’t let route calculations from grade school spoil the image of math for you!



FOUNDATIONS
FUNDAMENTAL RULES
MATHEMATICAL LOGIC
p ⇒ q
CONSISTENT SET OF AXIOMS?
NOPE
GÖDEL INCOMPLETENESS THEOREMS
SET THEORY
A
A∩B
B
X
f
Y
g∘f
g
Z
CATEGORY THEORY
THEORY OF COMPUTATION
0 0 0 1 1 B 0
q1
P≠NP?
COMPLEXITY THEORY
CARDINAL NUMBERS
ℵ0 ALEPH NULL
OCTONION
{e0,e1,e2,e3,e4,e5,e6,e7}
QUATERNION
a+bi+cj+dk
PI
π
EXPONENTIAL
e
COMPLEX NUMBERS
3, i, 4+3i, -4i
PRIME NUMBERS
3, 11, 47, 907
REAL NUMBERS
-4π, √2, e
RATIONAL NUMBERS
-7, 1/2, 2.32
INFINITY
∞
NUMBER THEORY
INTEGERS
...-2,-1,0,1,2...
NATURAL NUMBERS
1,2,3,4,5...
+ −
× ÷
ARITHMETIC
COMBINATORICS
PARTITION THEORY
TREE
GRAPH THEORY
GROUP THEORY
PERMUTATION GROUP
LINEAR ALGEBRA
NUMBER SYSTEMS
30
15
6
10
5
3
2
1
FACTORS OF 30
ORDER THEORY
MEASURE THEORY
MATRICES
ALGEBRA
x²-4x-8=5x+28
x²-9x-36=0
(x+3)(x-12)=0
EQUATION
y=mx+c
VECTORS
STRUCTURES
PURE MATHEMATICS
MÖBIUS STRIP
TOPOLOGY
DIFFERENTIAL GEOMETRY
1 HOLE (GENUS 1)
MANDELBROT SET
sin(θ)
cos(θ)
PYTHAGORAS
TRIGONOMETRY
GEOMETRY
SPACES
CHANGES
COMPLEX ANALYSIS
FRACTAL GEOMETRY
DYNAMICAL SYSTEMS
FLUID FLOW
BUTTERFLY EFFECT
CHAOS THEORY
ECOSYSTEMS
VECTOR CALCULUS
CALCULUS
DIFFERENTIAL
GRADIENT = dy/dx
f(x)
INTEGRAL
AREA = ∫₂⁹ f(x) dx
DIFFERENTIAL EQUATIONS
PERSIA c.820
ALGEBRA
c.1730 MATHEMATICAL NOTATION
INDIA c.628
FIRST ZERO
0
e^{iπ}=-1
COUNTING
NEGATIVE NUMBERS
-8
CHINA 200 BCE
50,000 BCE
GREECE 600-300 BCE
EGYPT
FIRST EQUATION 3000 BCE
ORIGINS
CRYPTOGRAPHY
ENCRYPT
DECRYPT
PUBLIC KEY
PRIVATE KEY
COMPUTER SCIENCE
TURING MACHINE
MACHINE LEARNING
CAT
while awake:
do_science()
if self.tired():
awake = False
self.repair_brain()
OPTIMIZATION
PROBABILITY
BAYES' RULE
P(A|B) = P(B|A)P(A) / P(B)
STATISTICS
MATHEMATICAL FINANCE
GAME THEORY
ECONOMICS
APPLIED MATHEMATICS
NUMERICAL ANALYSIS
ENGINEERING
CONTROLLER
SYSTEM
SENSOR
CONTROL THEORY
MATHEMATICAL PHYSICS
THEORETICAL PHYSICS
MATHEMATICAL CHEMISTRY
BIOMATHEMATICS
BY DOMINIC WALLIMAN © 2017
YOUTUBE: THE MAP OF MATHEMATICS

Life in general is the epitome of complex structure. It is only natural that mathematics be a core method of its investigation.

Applied Mathematics

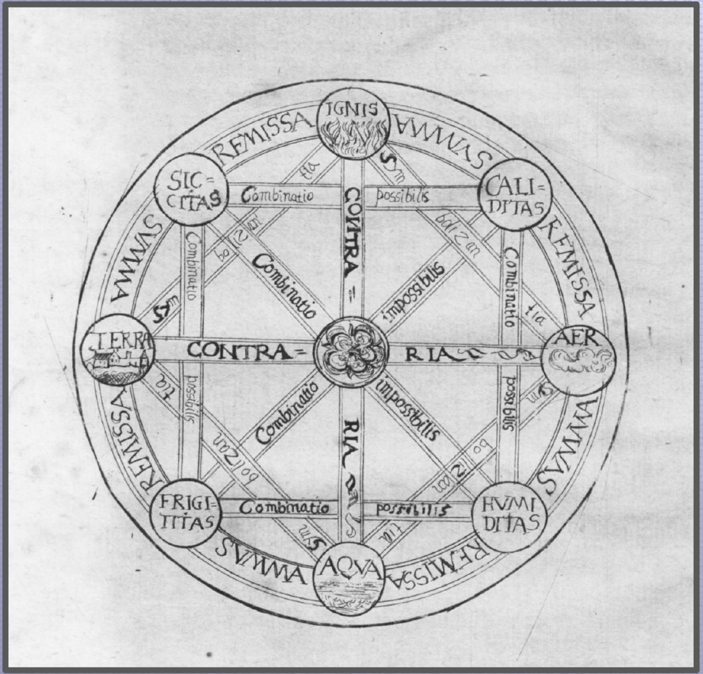


- Applied mathematics is concerned with finding parallels between the broad universe of mathematical abstractions and instances of those abstractions in the real world. This is a process called **Mathematical Modeling**.
- Artificial Intelligence is properly a topic of applied mathematics. We are interested in finding mathematical abstractions that capture the mechanics and nature of intelligence in the world.

Mathematics <--> Applied Mathematics

- In fields such as AI, we often find it difficult to find the proper mathematical model that captures the immense complexity under our consideration.
- When an application of mathematics is as demanding as AI is, there often needs to be work done in pure mathematics that can then be used as a model of real world phenomena.
- Neural Networks are good examples of mathematical structures that are well defined in terms of traditional mathematics, but do not have an overarching mathematical theory which allows us to prove properties of these objects. Proving properties of neural networks in general is pure mathematics, while figuring out how to use them in applications is applied mathematics.

AI as Logic, Language, and Symbolic Reasoning



[Characteristica Universalis]

- How do we find a set of axioms and rules of deduction that produce intelligent behavior?
- Language is an instance of logic. We have a set of concepts denoted by words, akin to axioms or fundamental units, and a way of combining these indivisible units via a grammar.
- How do we pose things like visual perception as logical/linguistic problems?

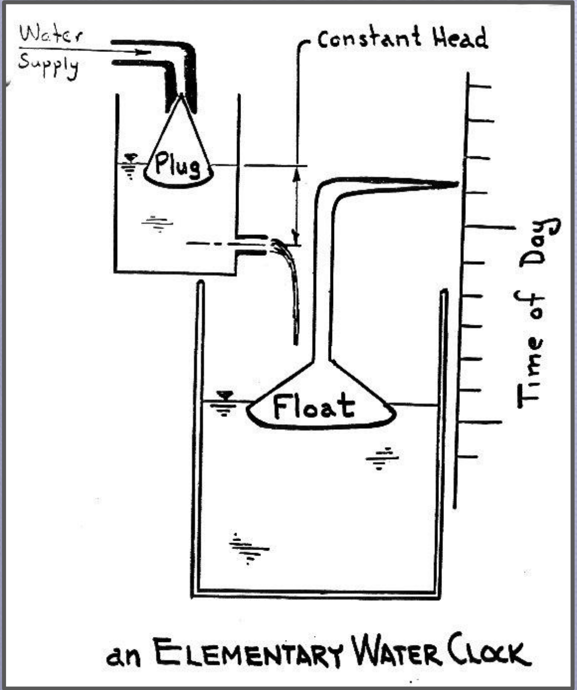
- Logic and language are forms of symbolic reasoning. Symbolic reasoning marked most early work on AI.

AI as Control/Cybernetics and Mechanics

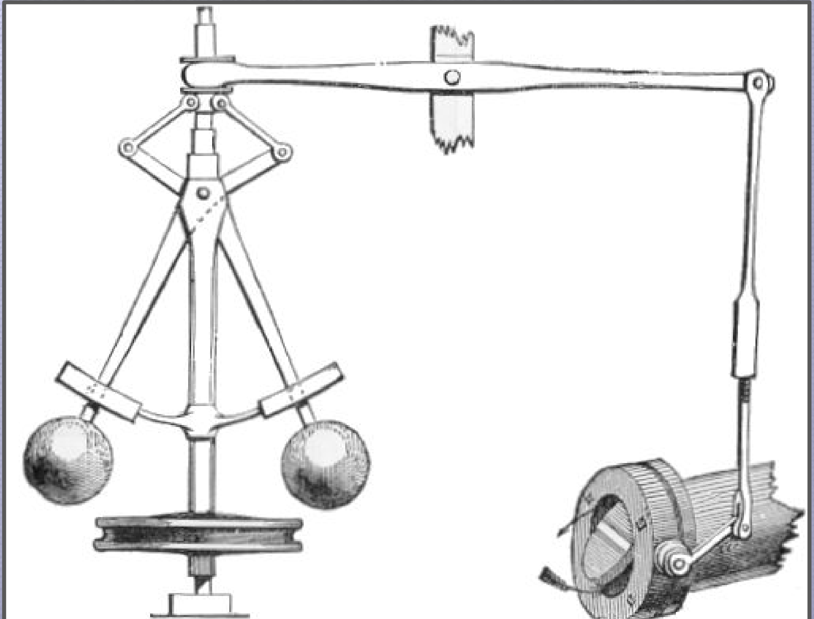
Why can't we derive intelligent machines from the first principles of physics?

- Control theory: "the control of continuously operating dynamical systems in engineered processes and machines. The objective is to develop a control model for controlling such systems using a control action in an optimum manner without delay or overshoot and ensuring control stability."
- Cybernetics: "the scientific study of control and communication in the animal and the machine." - Wiener
- Mechanical approaches to AI focus on things like homeostasis and equilibrium in respect to some set point.
- In respect to psychology, the mechanical approach has parallels to behaviorism. Cognitive psychology arose in response to this.
- The Little Writer.

Early Homeostatic Devices



16'th Century **BC**



Centrifugal Governor

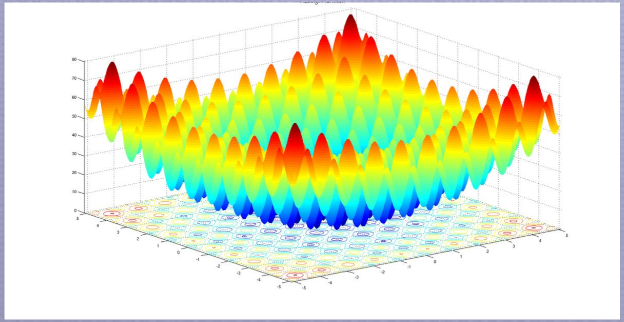
17'th Century **AD**

AI as Search

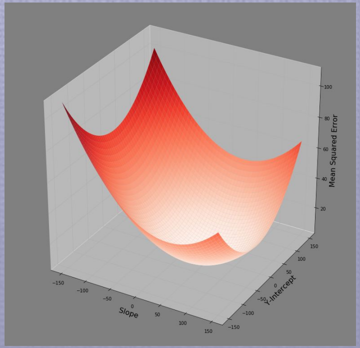
- If we had a computer that was infinitely powerful, all games, optimization, and well formed problems could be solved by brute force calculation.
  - The General Problem Solver (GPS) by Herbert A. Simon, J.C. Shaw, and Allen Newell in 1959 is an example of this. If one could sufficiently define a problem, it could in theory be solved by GPS.
- The problem is that we don’t have infinitely powerful computers. Thus, figuring out how to search explosive combinatorial spaces is fundamental to problem solving AI and the basis of knowledge in general.
- A* search, an extension of dijkstra's algorithm, is a famous example in early AI. It can markedly be used with definitions of heuristics to improve performance.

AI as Optimization

- You have many variables. The inner product of these variables produces a space. Certain regions of this space produce optimal solutions.
- We generally use calculus to navigate the curvature of these spaces.
- Higher dimensions, of problems of more than 3 variables, become hard to visualize.
- Sometimes the optimization space is hard to traverse for optimal solutions.



Rastrigin Function
The worst example of a hard to optimize space



Convex Optimization

Machine Learning

- Machine Learning is an approach to AI that poses intelligence and different computational tasks as a learning problem.
- There are three main areas of machine learning:
  - **Supervised Learning**: Learning from human labeled data. (Ex. Neural Networks)
  - **Unsupervised Learning**: Learning structure from data directly. (Ex. K-Means)
  - **Reinforcement Learning**: Learning from trial and error in an environment. (Q learning)
- There are many approaches to machine learning, among which are:
  - Neural Networks
  - Support Vector Machines
  - Genetic Algorithms: Algorithms that mimic biological evolution. Can be used with neural networks. In the early days of NN, before backprop, NN were trained by GA.
  - Statistical Methods: Bayesian Networks, Markov Methods,.. Etc.

Why Today's AI Starts with Data

- Only in recent times have we had the massive amounts of data and computer power that is available to us.
- Powerful supervised learning methods were developed in the 80’s onwards, but were not viable on the simple systems available then.
- Following the massive success of neural networks, people have turned to them for solving many different problems. Neural Networks also notoriously need lots of data.
- Different data captures different aspects of the world and lends itself to different forms of knowledge and analysis.
- Data science is the application of AI and statistical methods to studying data. Data science is highly practical and applicable to many business objectives.

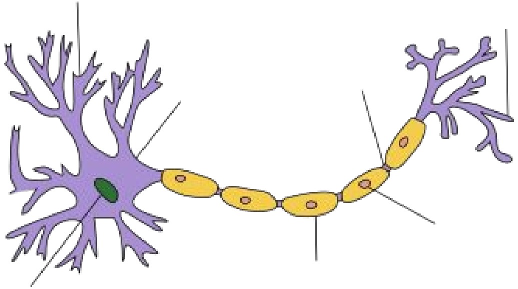
Neural Networks

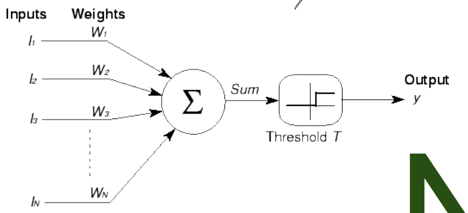
Neural Networks - Overview

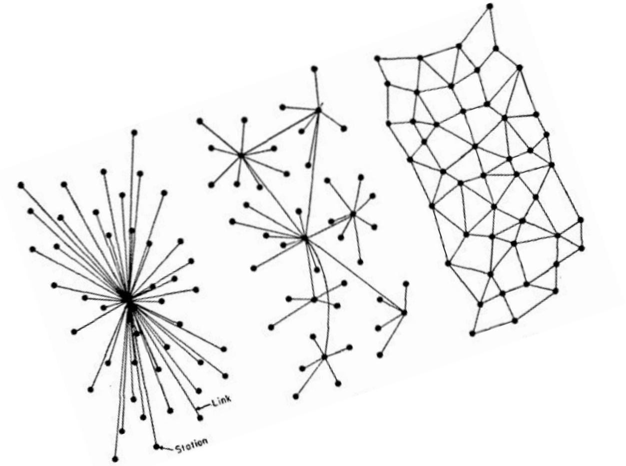
- Neurons
- Networks
- Neural Networks
- Connectionism
- PDP Models
- Origin of Neural Networks
- Components of a Neural Network
- Stages of Training a NN
- What does a NN do?
- Types of NN
- MNIST Dataset

- Hinton MNIST Demo
- Deep Learning
- Neural Network Playground

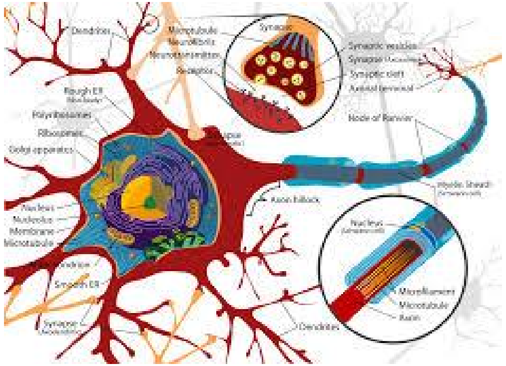
KEXP

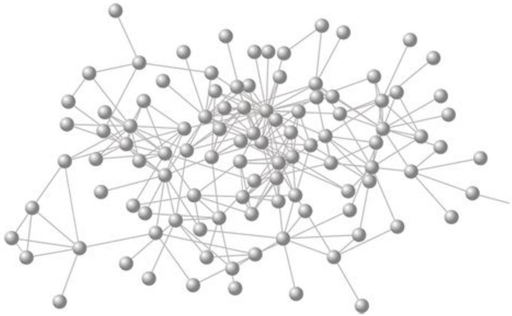






Neural Network

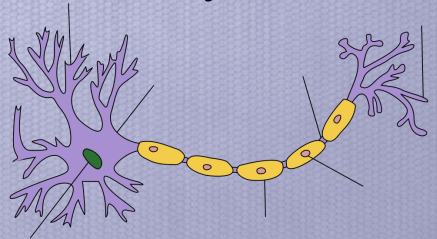




Neurons

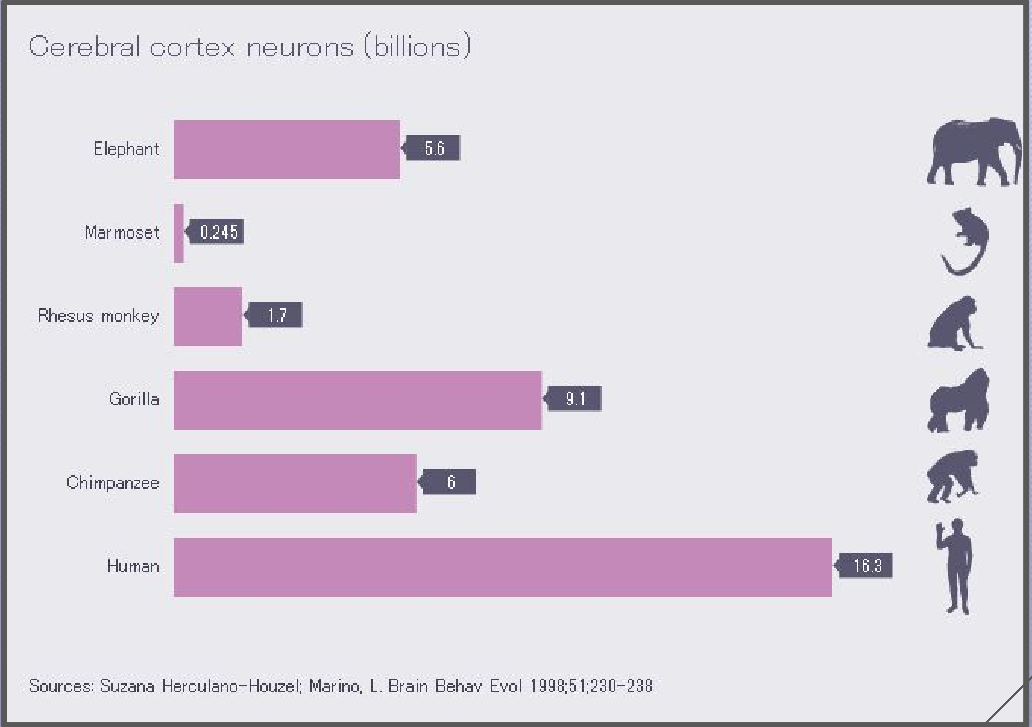
The brain has about 86 billion of these.

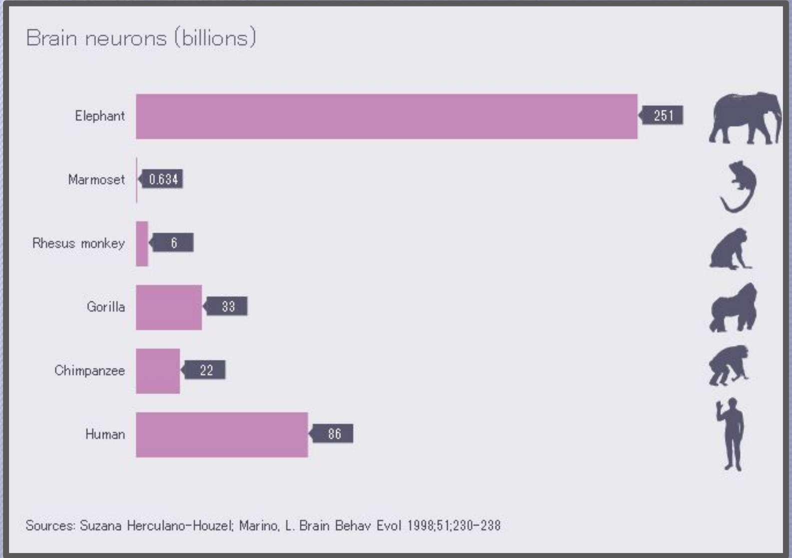
- Biological neurons are the core building block of the nervous system. Our nervous system is responsible for coordinating all the signals that keep different parts of our body working together as a whole.
- Neurons become specialized for different tasks. (i.e. sensory and motor)
- Neurons have three main components:
  - The Cell Body (soma)
  - Axon - Sending Signals
  - Dendrites - Receiving Signals

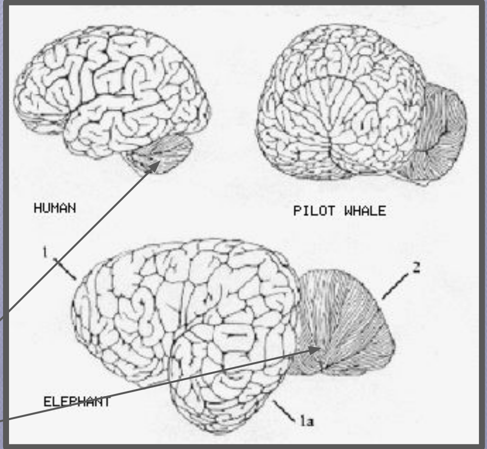


- Neurons communicate with chemicals called Neurotransmitters and by electrical signals.
- Artificial Neurons are a mathematical model of the basic computational functions of a neuron. They do not come close to the true complexity of biological neurons; however, they are still amazingly useful.

Neurons in Animals







SOURCE

SOURCE

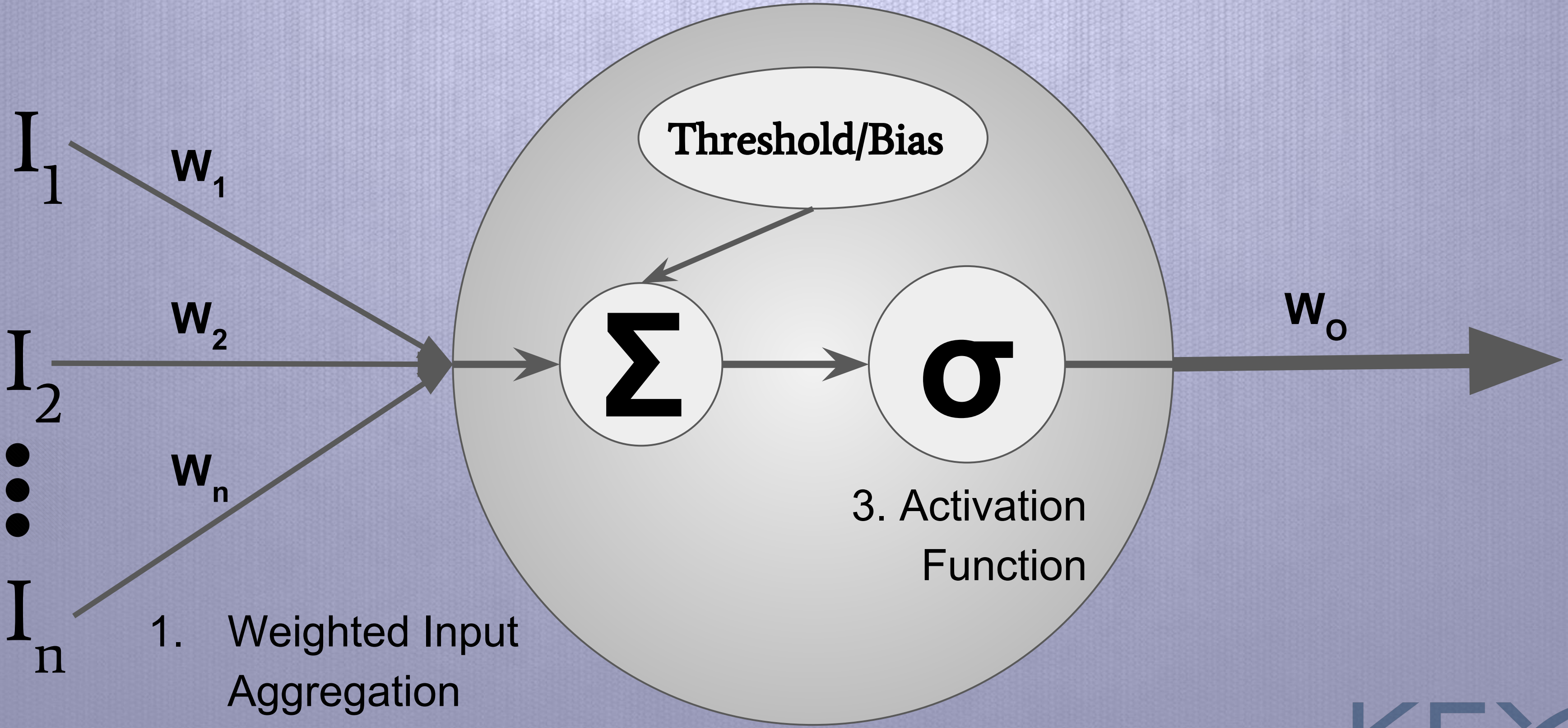
Why aren't elephants running the world?

Cerebrum vs Cerebellum

Artificial Neurons

In linear algebra:
$\sigma(I^TW + B)$

2. Threshold/Bias

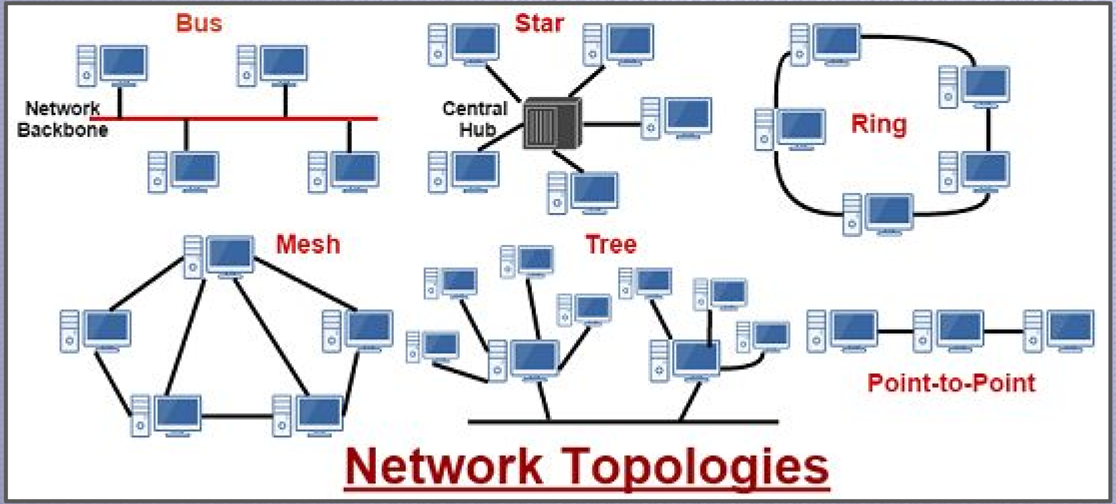


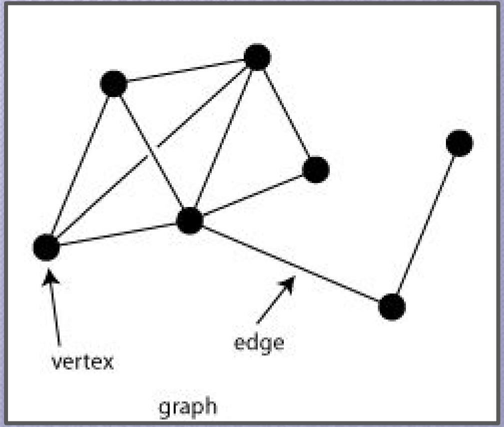
3. Activation Function

1. Weighted Input Aggregation

KEXP

Networks

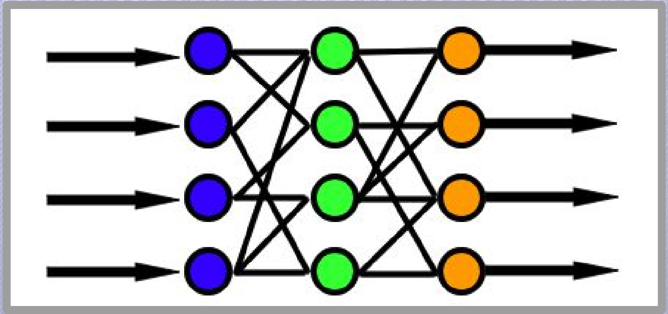




Example: computer networks

- Networks describe the connectivity of distinct things.
  - Nodes: The entities in the network. (Neurons)
  - Edges: The connections between the entities (Axons)
- Networks are modeled by mathematical structures called graphs.

Neural Networks



- Neural Networks are networks of neurons.
- These networks can be trained on data to produce useful representations that enable accurate predictions.
- Neurons become specialized during training. And groups of neurons take on a shared representation (localization).
- Neural networks provide a link between optimization and knowledge representation.

Connectionism

- “Connectionism within cognitive science is a theory of information processing. Unlike classical systems which use explicit, often logical, rules arranged in an hierarchy to manipulate symbols in a serial manner, however, connectionist systems rely on parallel processing of sub-symbols, using statistical properties instead of logical rules to transform information. Connectionists base their models upon the known neurophysiology of the brain and attempt to incorporate those functional properties thought to be required for cognition.” - Medler
- In this view, intelligence emerges from a complex network of simple units.
- In the brain, where does in intelligence start? Is a neuron intelligent? Is a group of neurons intelligent?

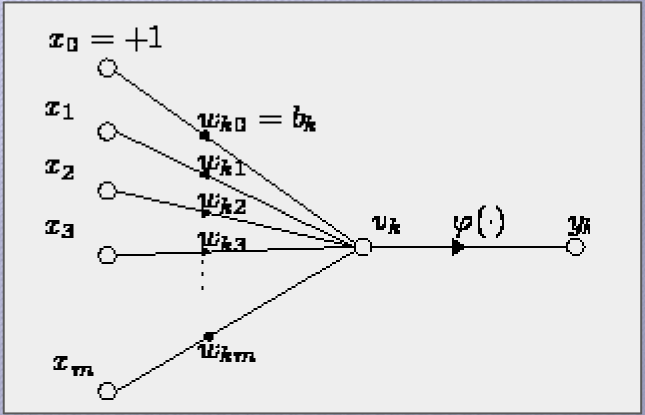
Parallel Distributed Processing (PDP)

1. A set of processing units.
2. A state of activation.
3. An output function for each unit.
4. A pattern of connectivity among units.
5. A propagation rule for propagating patterns of activities through the network of connectivities.
6. An activation rule for combining the inputs impinging on a unit with the current state of that unit to produce a new level of activation for the unit.
7. A learning rule whereby patterns of connectivity are modified by experience.
8. An environment within which the system must operate.

Rumelhart, Hinton, and McClelland

"Either the universe is composable or God exists."
-I heard Yann LeCun paraphrase this quote

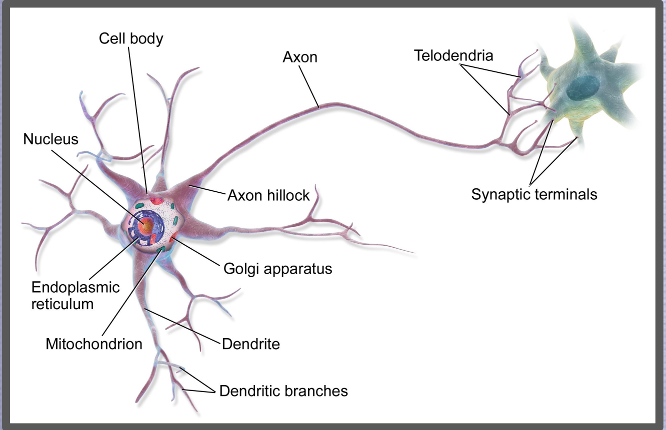
Origin of Neural Networks



The Modern Connectionist Timeline

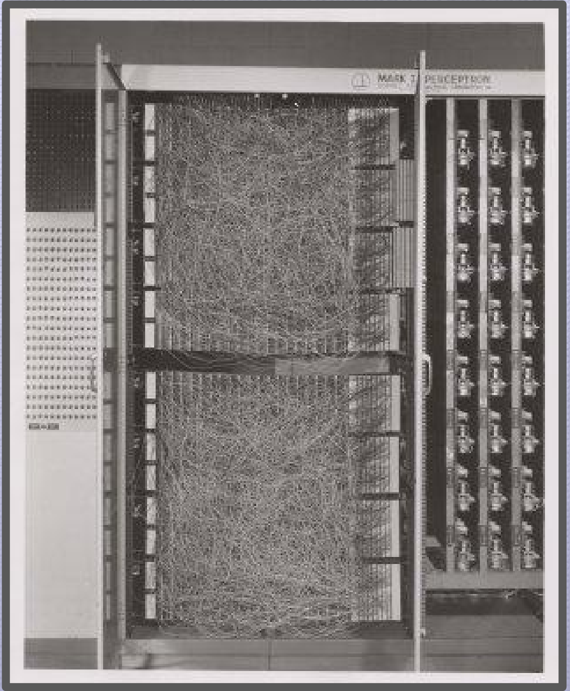
- 1943: Threshold Logic (McCulloch and Pitts)
- 1954: Hebbian Networks (Wesley A. Clark)
- 1958: Perceptrons (Frank Rosenblatt)
- 1969: AI Winter (Minsky and Papert)
- 1974: Multi-Layer Perceptrons and Backpropagation (Werbos)
- 1990: Convolutional Neural Networks (LeCun first runaway success)
- 1997: Long Short-Term Memory Networks (Hochreiter & Schmidhuber)
- 2014: Generative Adversarial Networks (Goodfellow et al.)

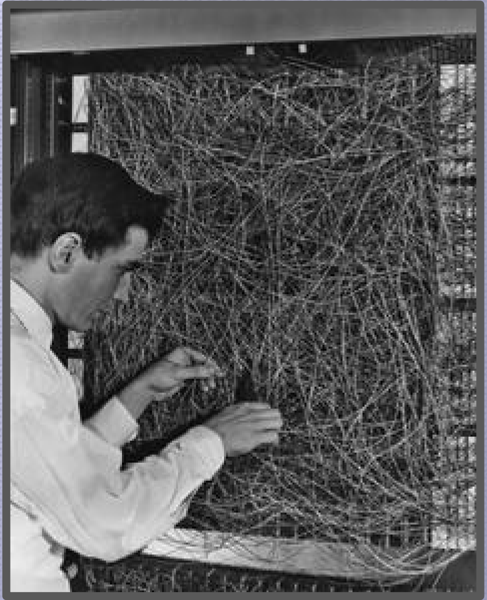
***An incomplete history**



The Neuron

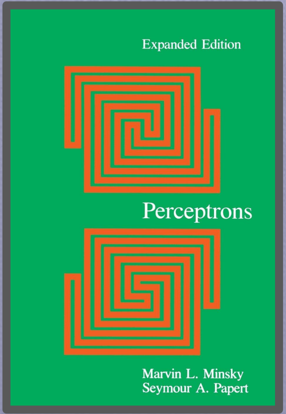
Biological Inspiration





Mark 1 Perceptron

Frank Rosenblatt



Harbingers of the AI Winter

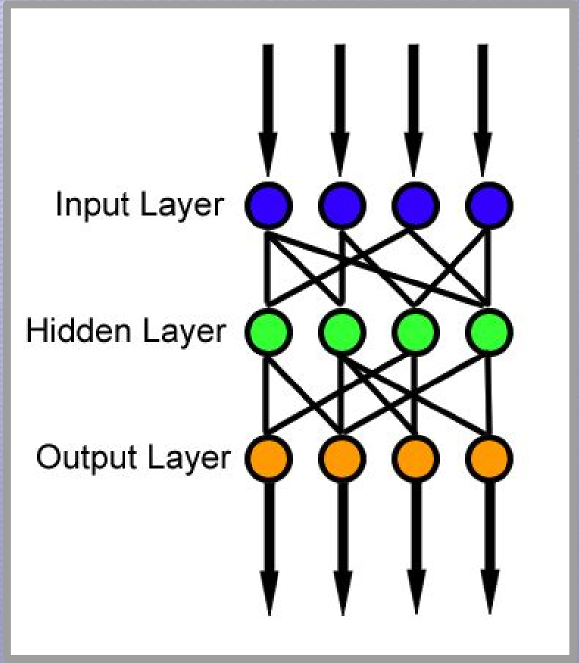
KEXP

Components of a Neural Network

- **The Objective/Loss function**
    - Given the last layer of a neural network and some targets for the output of the network, how well did the network perform?
- **The Network Gradient**
    - Backpropagation of errors from the cost function backwards through the network. This is really the chain rule of calculus.
- **The Optimization Algorithm**
    - Gradient Descent iteratively uses the network gradient, computed by backpropagation, to adjust the weights of the network. Adjusting these weights is called training or learning.
- **The Network Architecture**
    - How many neurons and layers a network has. How the neurons/layers are connected (Feedforward, Recurrent, Etc). What types of activations functions the neurons have.

Stages of Training Neural Network

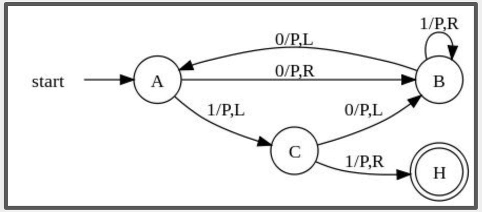
1. Data Preparation and Initialization
2. Training Loop - Training Set
    a. Forward Propagation of Training Batch
    b. Backpropagation of Errors on Training Targets
3. Evaluation
    a. Test Set
        i. Generalization of Training to the Test Set
    b. Validation Set
        i. Hyperparameter Optimization



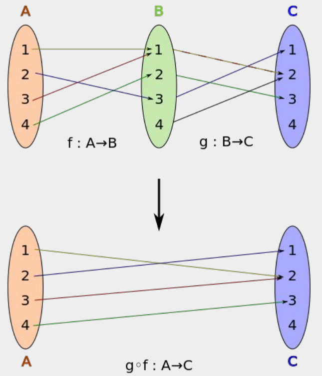
What does a neural network do? Two Perspectives

- Turing Machines and the Lambda Calculus are both universal models of computation.
- A universal model of computation is one in which any computable function can be represented (see the Church-Turing Thesis)
- Turing Machines pose computations as state based programs.
- The Lambda Calculus poses computation as recursive functions.

Neural Networks learn programs



Neural Networks learn functions.



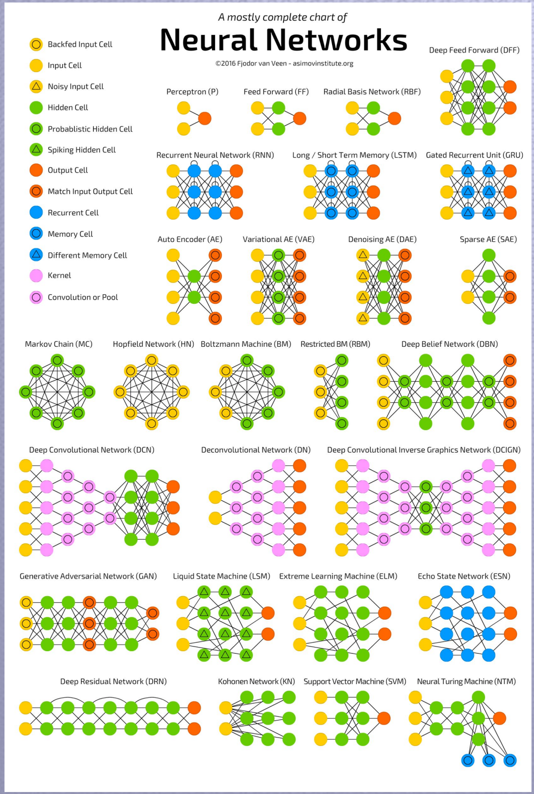
Function Approximation- Non-Recursive Functions

- Non-Recursive Functions have no feedback loop. They are particular mapping of one set to another, from one space to another.
- Linear functions can be defined by matrices and can be thought as a linear transformation of space.
- Non-Linear functions are more complex.
- All non-novel neural networks approximate nonlinear functions.
- Activation functions of a neural network introduce nonlinearity.

Program Approximation - Recursive Functions

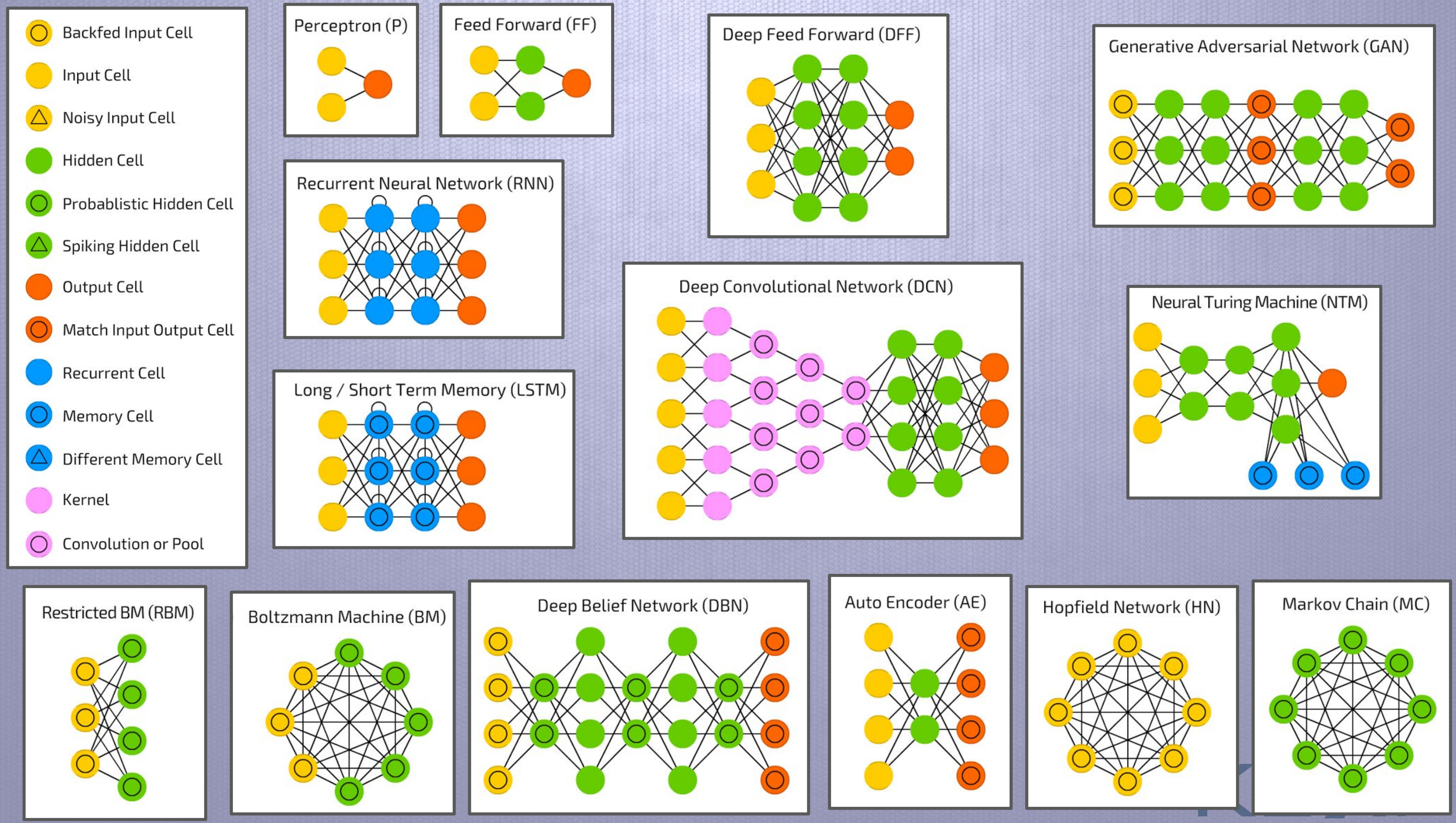
- The Church-Turing Thesis tells us that turing machines are mathematically isomorphic to a particular set of recursive functions called general recursive functions -- defined by Godel in 1933.
- A recurrent neural network is capable of approximating general recursive functions.
- Programs = Recursive Functions

Types of Neural Networks



The number of neural network is only limited by our creativity.

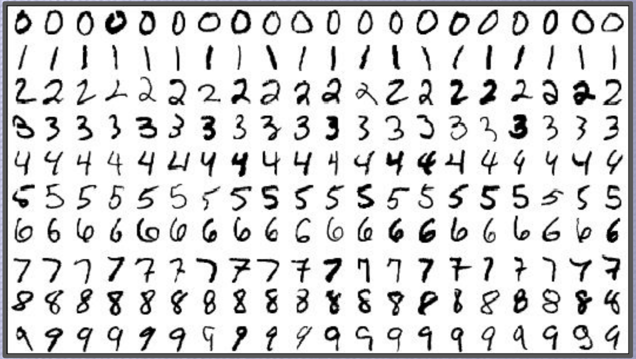
Some have well known properties that make them useful for different tasks.



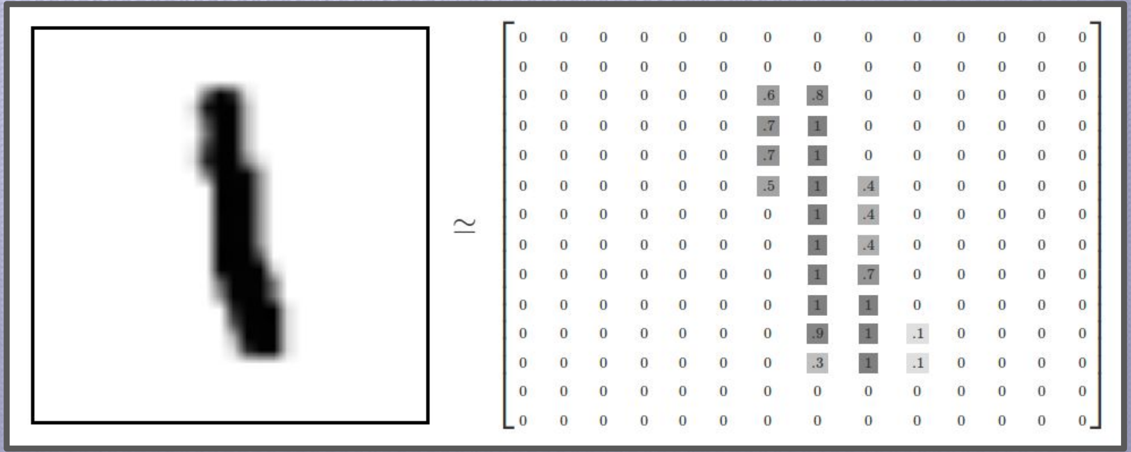
Backfed Input Cell
Input Cell
Noisy Input Cell
Hidden Cell
Probablistic Hidden Cell
Spiking Hidden Cell
Output Cell
Match Input Output Cell
Recurrent Cell
Memory Cell
Different Memory Cell
Kernel
Convolution or Pool
Perceptron (P)
Feed Forward (FF)
Deep Feed Forward (DFF)
Generative Adversarial Network (GAN)
Recurrent Neural Network (RNN)
Long / Short Term Memory (LSTM)
Deep Convolutional Network (DCN)
Neural Turing Machine (NTM)
Restricted BM (RBM)
Boltzmann Machine (BM)
Deep Belief Network (DBN)
Auto Encoder (AE)
Hopfield Network (HN)
Markov Chain (MC)

MNIST Dataset [1]

55,000 28 X 28 grayscale images of numerical digits [0,9], each with a class label.

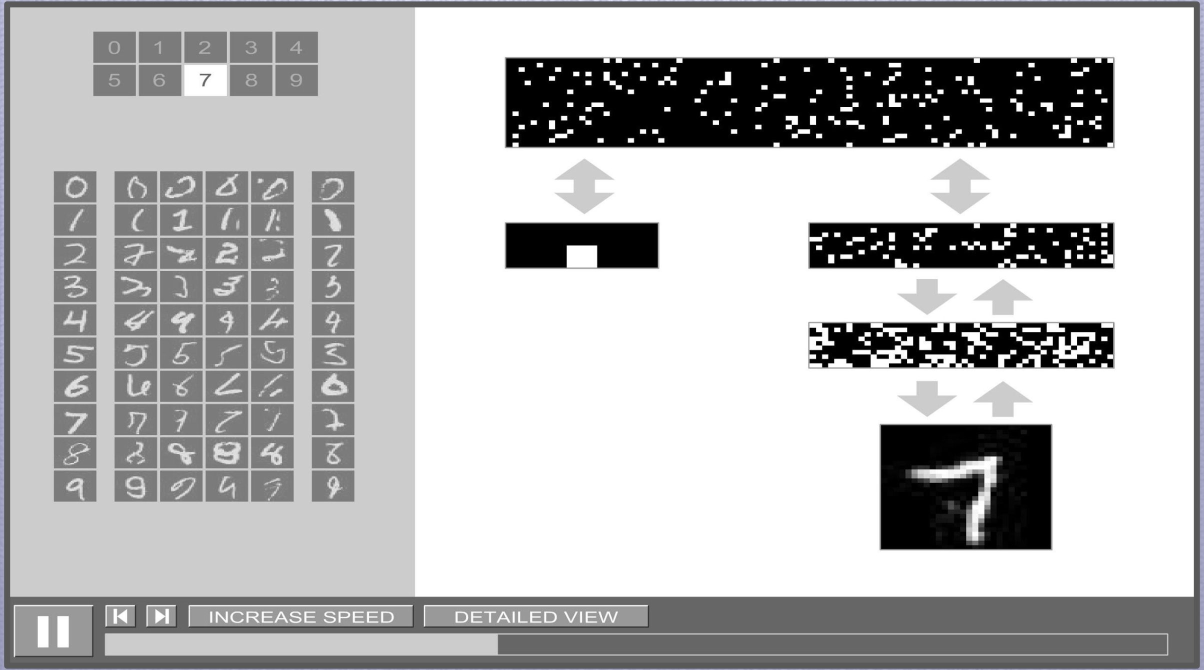


The images are unrolled into a column vector of size:
28 X 28 = 784
So the dimension of the data matrix is:
55,000 X 784
And the label vector is:
55,000 X 1



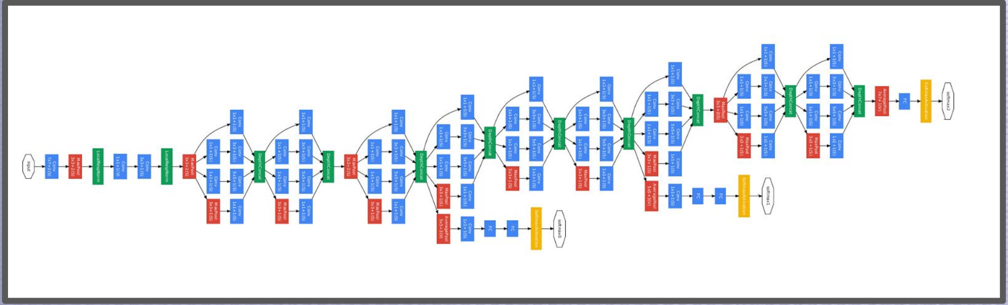
- Often referred to as the Fruit Fly of machine learning. It is a dataset that is simple, but sufficiently complex to showcase different algorithms.

Hinton MNIST Demo

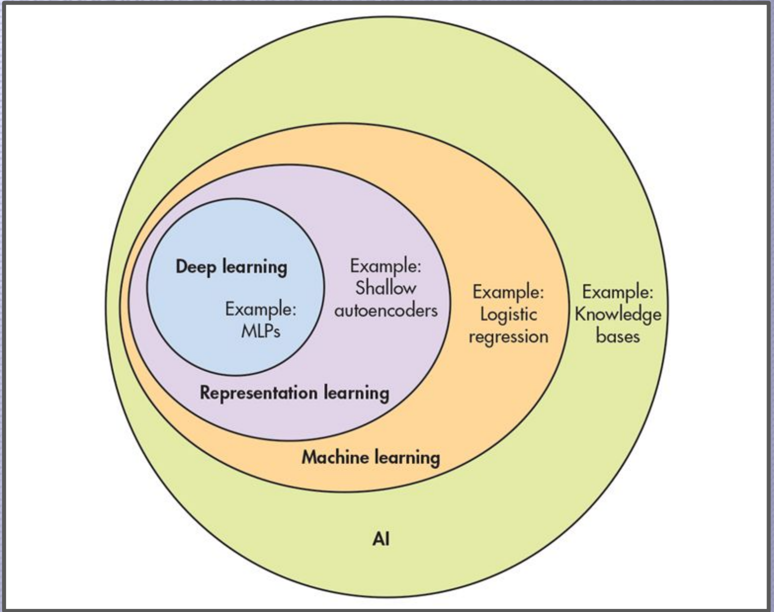


- http://www.cs.toronto.edu/~hinton/adi/index.htm

Deep Learning

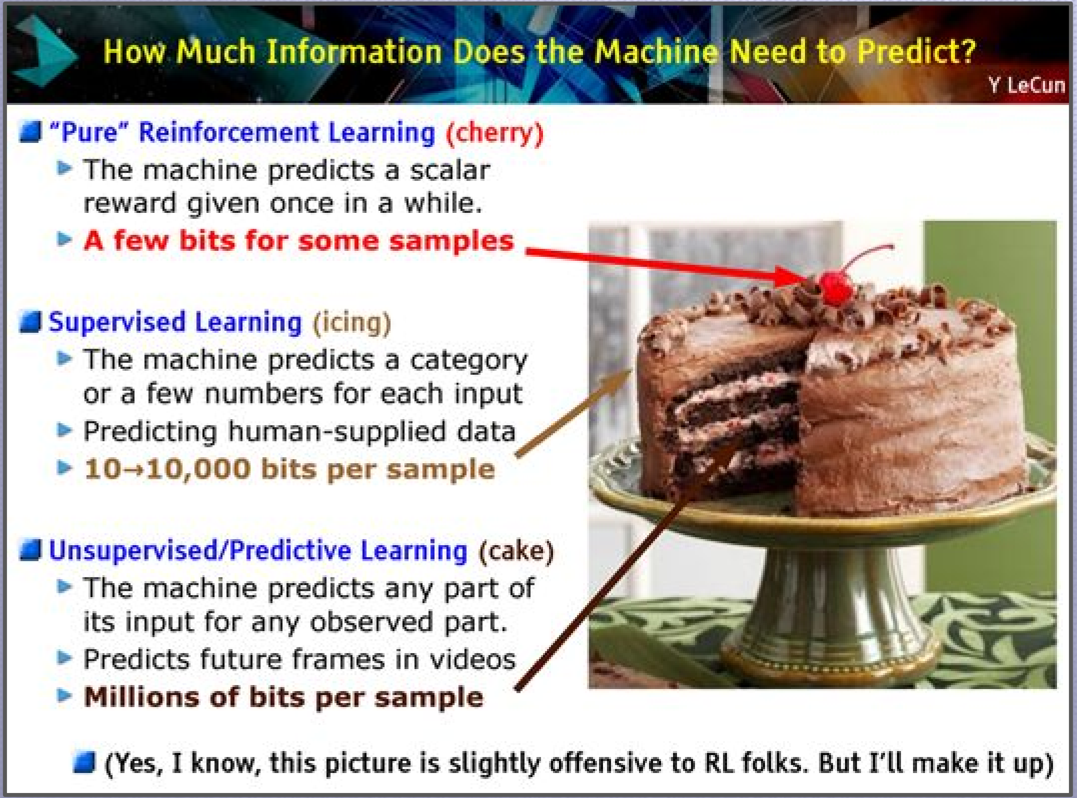


GoogleNet



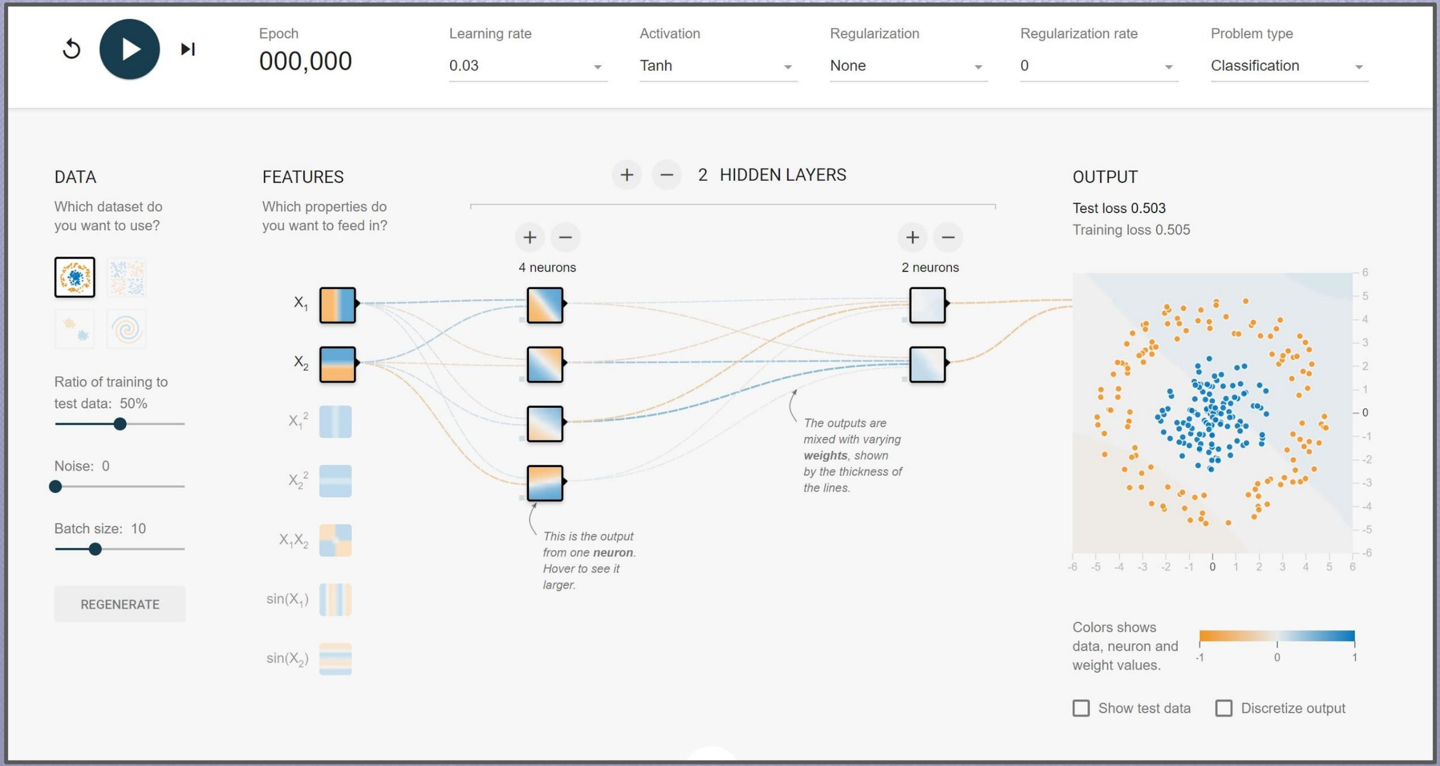
- Deep Learning refers to the observation that neural networks with many layers -- this is the deep part,-- are easier to train and generalize better than shallow neural networks.
- The term is also used to encapsulate modern neural network research.

Deep Learning - The Cake



KEXP

Neural Network Playground



http://playground.tensorflow.org

KEXP

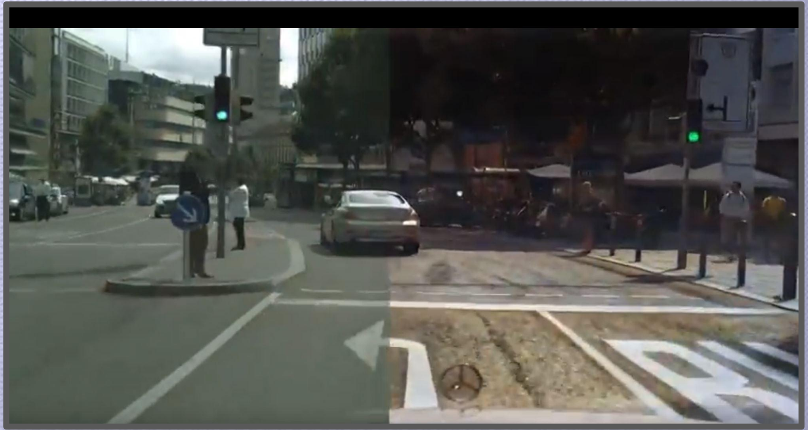
DISCLAIMER

Neural Networks are not the only approach to modern machine learning! And in some cases, they are not the most effective nor practical. If the only tool you have is a hammer, it is tempting to see all problems as a nail.

Deep Learning Applications

Image-to-Image Translation - GANs

VIDEO



VIDEO

Video

VIDEO

- CycleGan - https://junyanz.github.io/CycleGAN/
- Pix2Pix Demo: https://affinelayer.com/pixsrv/
- Pix2Pix HD: https://tcwang0509.github.io/pix2pixHD/

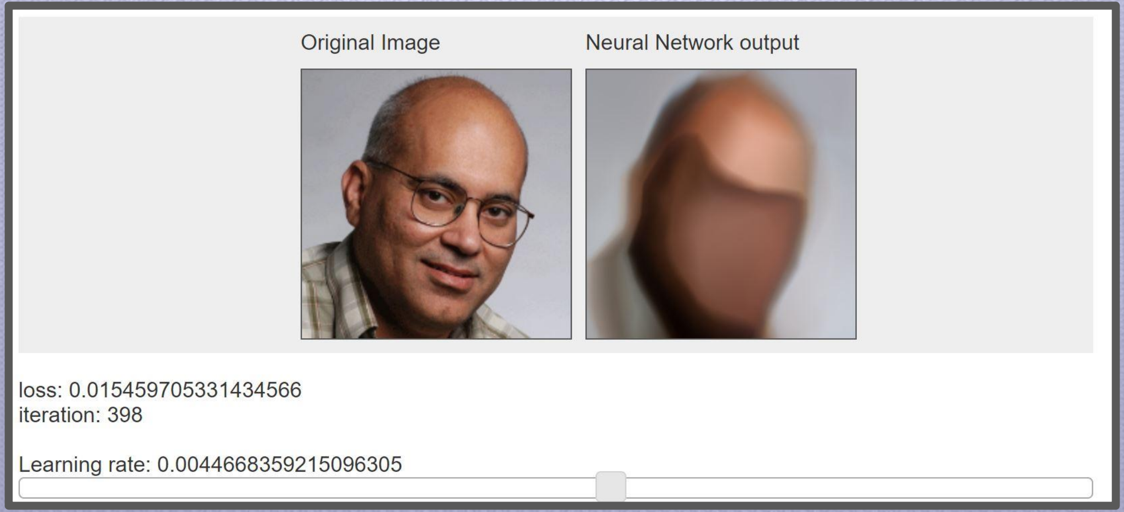
Image Synthesis - GANs

Video | Repo

Video

- https://www.csail.mit.edu/news/creating-videos-future
- http://carlvondrick.com/tinyvideo/

Image Painting - Autoencoders



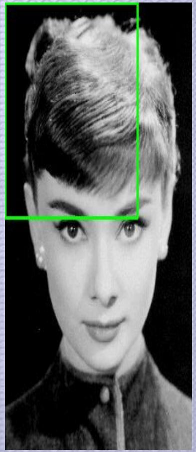
- Input: (x,y) position. Output: (r,g,b) color value.
- The network tries to learn a neural representation that produces the original image. This is similar in nature to an autoencoder.

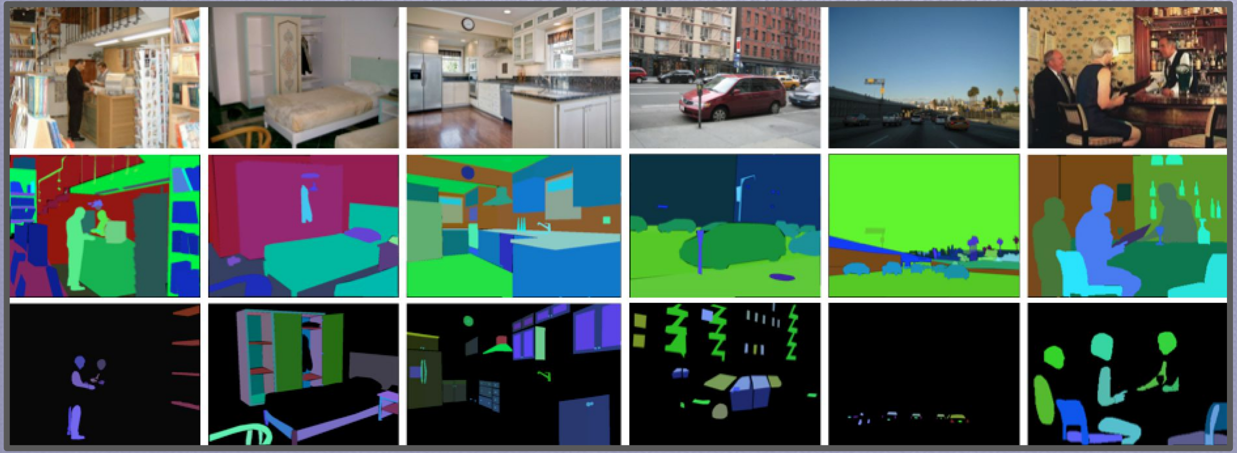
Source: https://cs.stanford.edu/people/karpathy/convnetjs/demo/image_regression.html

KEXP

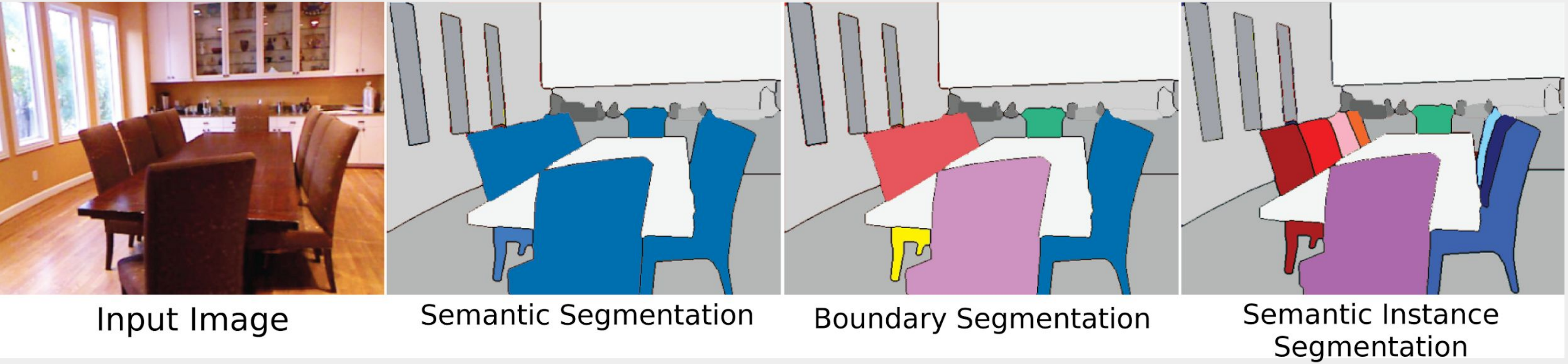
Segmentation - CNNs

- Semantic Segmentation - Pixel-Wise Labeling
- Boundary Segmentation - Cohesive Regions
- Instance Segmentation - Countable Regions





Object classes: chair, table, window, etc...



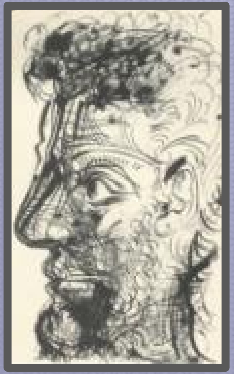
Source: http://cs.nyu.edu/~silberman/images/coverage.png

Style Transfer - CNNs

A Neural Algorithm of Artistic Style - https://arxiv.org/abs/1508.06576

content loss: 1.22706e+06
style loss: .659507
total loss: 1.89246e+06

Style



Content

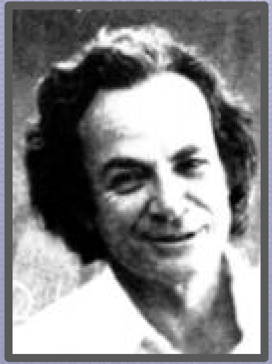
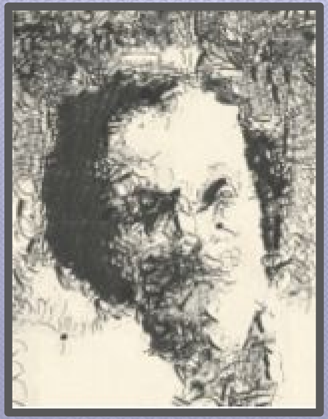
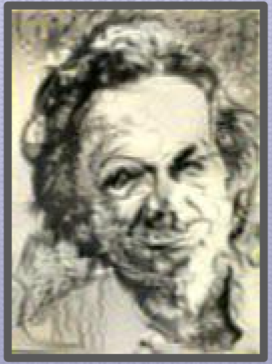
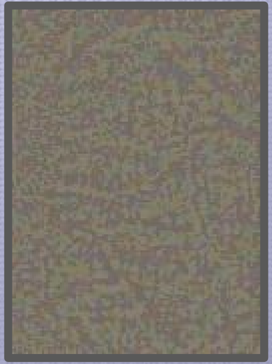


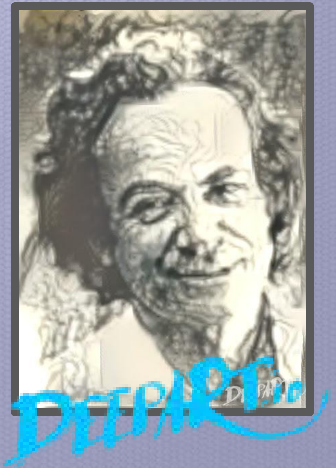
Image Quilting



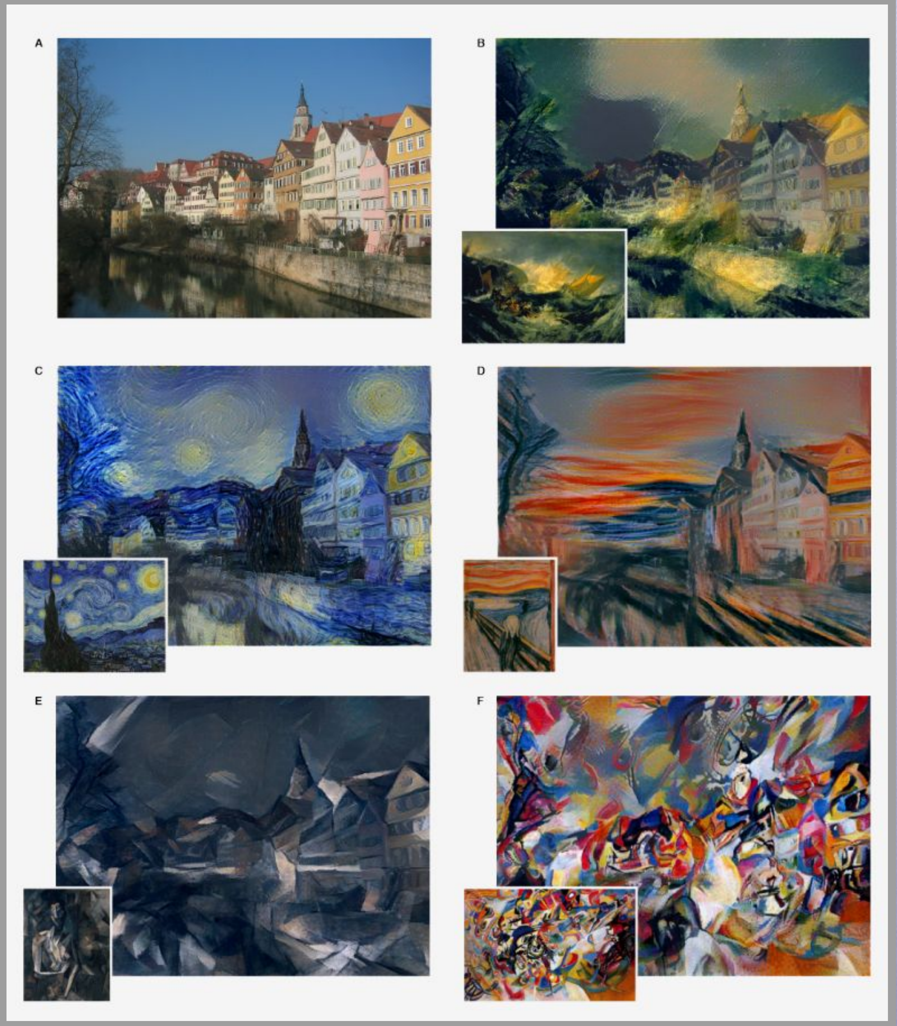
Neural



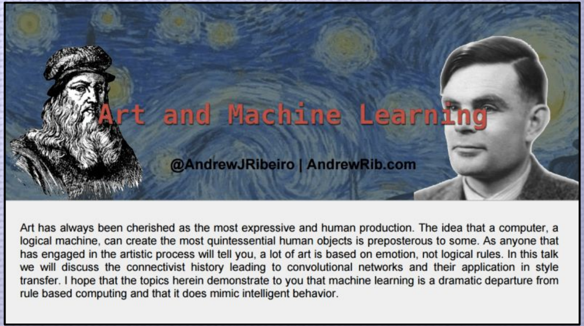




Note: I took a screenshot from the paper to get the style,content, and image quilting result images. They were probably scaled down in the paper so we didn't get great results, but it's still illustrative.



[DEMO]



[Slides]

[https://www.youtube.com/watch?v=Khuj4ASldmU]

Sequence Modeling: Recurrent Neural Networks

- Handwriting Generation
- Machine Translation and Language Modeling
- Autocomplete
- Automated Image Captioning W/ Attention Modeling
- Audio Synthesis (Text-to-speech)
- Voice Recognition
- Conversational chat-bots

Handwriting Generation

https://distill.pub/2016/handwriting/

Audio Synthesis - Wavenet

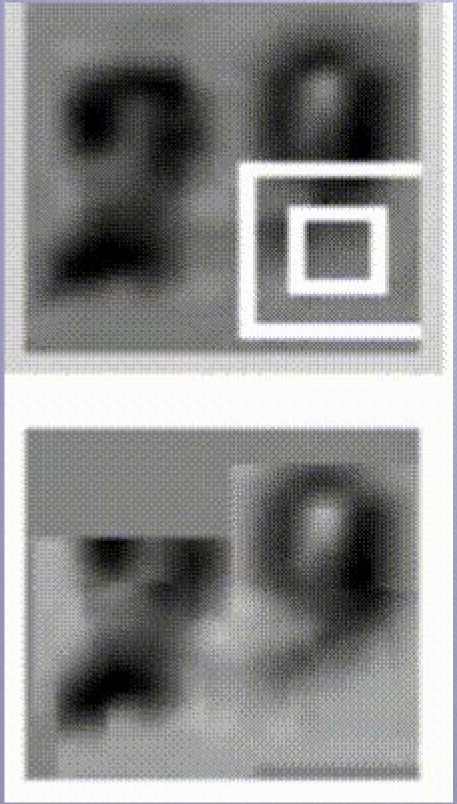
WaveNet: A Generative Model for Raw Audio

This post presents **WaveNet**, a deep generative model of raw audio waveforms. We show that WaveNets are able to generate speech which mimics any human voice and which sounds more natural than the best existing Text-to-Speech systems, reducing the gap with human performance by over 50%.

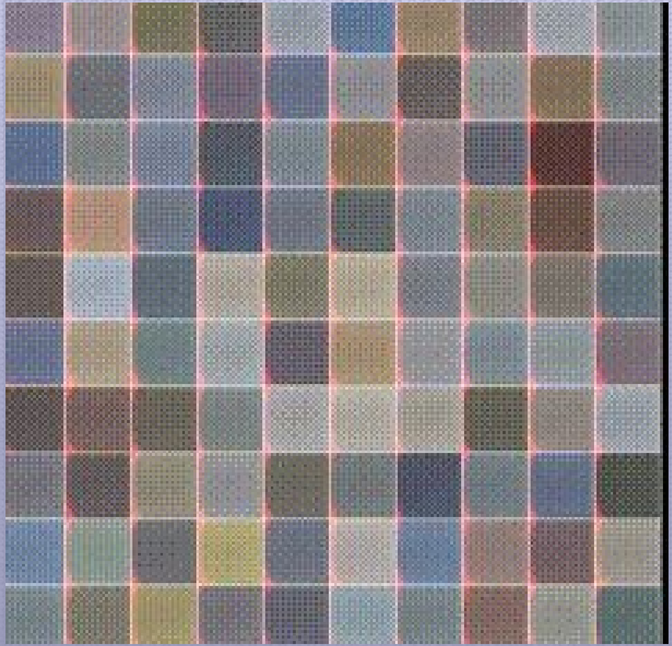
We also demonstrate that the same network can be used to synthesize other audio signals such as music, and present some striking samples of automatically generated piano pieces.

https://deepmind.com/blog/wavenet-generative-model-raw-audio/

Applications of RNN: Visual Attention



RNN learns to read house numbers.



RNN learns to paint house numbers.

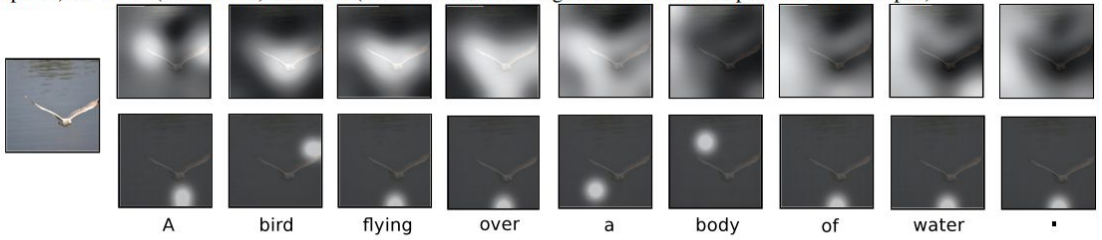
- The way humans perceive an image can be considered a time series problem; namely, where our focus changes over time.
- There has been research done on connecting RNN and CNN for image captioning. The idea here is that an RNN figures out the attention pattern and a CNN figures out what the objects are. This is referred to as an ensemble method.
- Attention is considered to be one of the big uses of RNN.

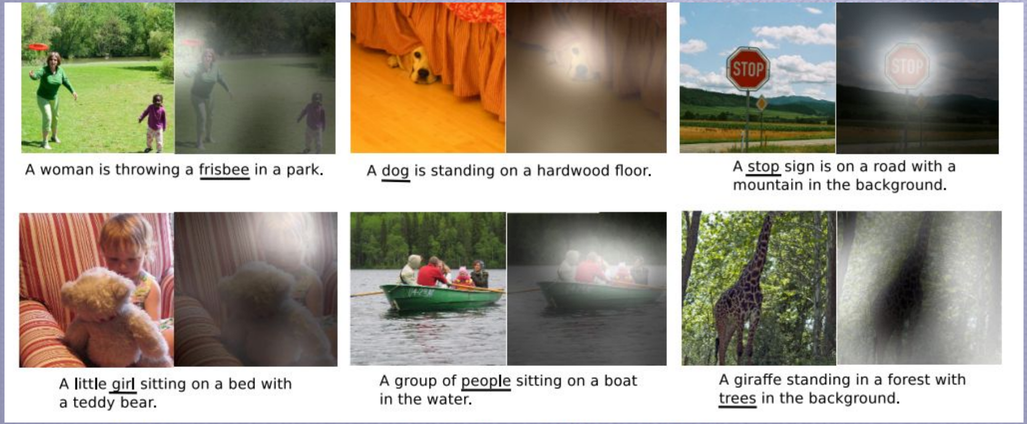
Source.

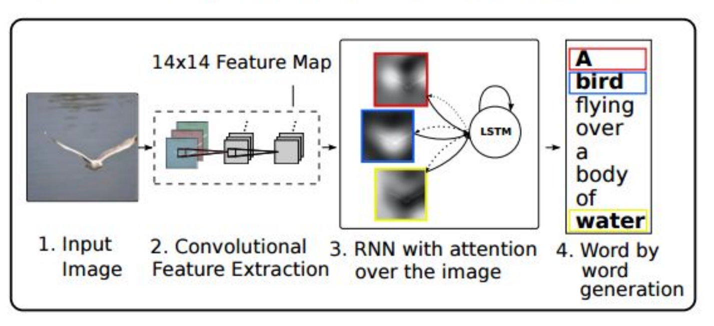
KEXP

Applications of RNN: Image Captioning

Figure 2. Attention over time. As the model generates each word, its attention changes to reflect the relevant parts of the image. "soft" (top row) vs "hard" (bottom row) attention. (Note that both models generated the same captions in this example.)







Source

Applications of RNN: Language Modeling

```
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:
Well, your wit is in the care of side and that.

Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:
Come, sir, I will make did behold your worship.

VIOLA:
I'll drink it.
```

Shakespeare

For $\bigoplus_{n=1,\ldots,m}$ where $\mathcal{L}_{m_\bullet} = 0$, hence we can find a closed subset $\mathcal{H}$ in $\mathcal{H}$ and any sets $\mathcal{F}$ on X, U is a closed immersion of S, then $U \to T$ is a separated algebraic space.

Proof. Proof of (1). It also start we get

$$S = \mathrm{Spec}(R) = U \times_X U \times_X U$$

and the comparicoly in the fibre product covering we have to prove the lemma generated by $\coprod Z \times_U U \to V$. Consider the maps M along the set of points Sch_{fppf} and $U \to U$ is the fibre category of S in U in Section, ?? and the fact that any U affine, see Morphisms, Lemma ??. Hence we obtain a scheme S and any open subset $W \subset U$ in $Sh(G)$ such that $\mathrm{Spec}(R') \to S$ is smooth or an

$$U = \bigcup U_i \times_{S_i} U_i$$

which has a nonzero morphism we may assume that f_i is of finite presentation over S. We claim that $\mathcal{O}_{X,x}$ is a scheme where $x, x', s'' \in S'$ such that $\mathcal{O}_{X,x'} \to \mathcal{O}'_{X',x'}$ is separated. By Algebra, Lemma ?? we can define a map of complexes $\mathrm{GL}_{S'}(x'/S'')$ and we win. □

To prove study we see that $\mathcal{F}|_U$ is a covering of $\mathcal{X}'$, and $\mathcal{T}_i$ is an object of $\mathcal{F}_{X/S}$ for $i > 0$ and $\mathcal{F}_p$ exists and let $\mathcal{F}_i$ be a presheaf of $\mathcal{O}_X$-modules on $\mathcal{C}$ as a $\mathcal{F}$-module. In particular $\mathcal{F} = U/\mathcal{F}$ we have to show that

$$\widetilde{M}^\bullet = \mathcal{I}^\bullet \otimes_{\mathrm{Spec}(k)} \mathcal{O}_{S,s} - i_X^{-1}\mathcal{F})$$

is a unique morphism of algebraic stacks. Note that

$$\mathrm{Arrows} = (Sch/S)^{opp}_{fppf}, (Sch/S)_{fppf}$$

and

$$V = \Gamma(S, \mathcal{O}) \longmapsto (U, \mathrm{Spec}(A))$$

is an open subset of X. Thus U is affine. This is a continuous map of X is the inverse, the groupoid scheme S.

Proof. See discussion of sheaves of sets. □

The result for prove any open covering follows from the less of Example ??. It may replace S by $X_{spaces,étale}$ which gives an open subspace of X and T equal to S_{Zar}, see Descent, Lemma ??. Namely, by Lemma ?? we see that R is geometrically regular over S.

Lemma 0.1. *Assume (3) and (3) by the construction in the description.*

Suppose $X = \lim |X|$ (by the formal open covering X and a single map $\underline{Proj}_X(\mathcal{A}) = \mathrm{Spec}(B)$ over U compatible with the complex

$$Set(\mathcal{A}) = \Gamma(X, \mathcal{O}_{X,\mathcal{O}_X}).$$

When in this case of to show that $\mathcal{Q} \to \mathcal{C}_{Z/X}$ is stable under the following result in the second conditions of (1), and (3). This finishes the proof. By Definition ?? (without element is when the closed subschemes are catenary. If T is surjective we may assume that T is connected with residue fields of S. Moreover there exists a closed subspace $Z \subset X$ of X where U in X' is proper (some defining as a closed subset of the uniqueness it suffices to check the fact that the following theorem

(1) *f is locally of finite type. Since $S = \mathrm{Spec}(R)$ and $Y = \mathrm{Spec}(R)$.*

Proof. This is form all sheaves of sheaves on X. But given a scheme U and a surjective étale morphism $U \to X$. Let $U \cap U = \coprod_{i=1,\ldots,n} U_i$ be the scheme X over S at the schemes $X_i \to X$ and $U = \lim_i X_i$. □

The following lemma surjective restrocomposes of this implies that $\mathcal{F}_{x_0} = \mathcal{F}_{x_0} = \mathcal{F}_{\mathcal{X},\ldots,0}$.

Lemma 0.2. *Let X be a locally Noetherian scheme over S, $E = \mathcal{F}_{X/S}$. Set $\mathcal{I} = \mathcal{J}_1 \subset \mathcal{I}'_n$. Since $\mathcal{I}^n \subset \mathcal{I}^n$ are nonzero over $i_0 \leq \mathfrak{p}$ is a subset of $\mathcal{J}_{n,0} \circ \overline{A}_2$ works.*

Lemma 0.3. *In Situation ??. Hence we may assume $\mathfrak{q}' = 0$.*

Proof. We will use the property we see that $\mathfrak{p}$ is the mext functor (??). On the other hand, by Lemma ?? we see that

$$D(\mathcal{O}_{X'}) = \mathcal{O}_X(D)$$

where K is an F-algebra where δ_{n+1} is a scheme over S. □

Latex

Source

KEXP

RNN Shakespeare

- 3-layer RNN with 512 hidden nodes on each layer.
- Trained on all the works of Shakespeare concatenated into a single (4.4MB) file.
- Source: http://karpathy.github.io/2015/05/21/rnn-effectiveness/

```
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:
Well, your wit is in the care of side and that.

Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:
Come, sir, I will make did behold your worship.

VIOLA:
I'll drink it.
```

Harry Potter: Written by Artificial Intelligence

LSTM Trained on First Four Harry Potter Books

“The Malfoys!” said Hermione.

Harry was watching him. He looked like Madame Maxime. When she strode up the wrong staircase to visit himself.

“I’m afraid I’ve definitely been suspended from power, no chance—indeed?” said Snape. He put his head back behind them and read groups as they crossed a corner and fluttered down onto their ink lamp, and picked up his spoon. The doorbell rang. It was a lot cleaner down in London.

Hermione yelled. The party must be thrown by Krum, of course.

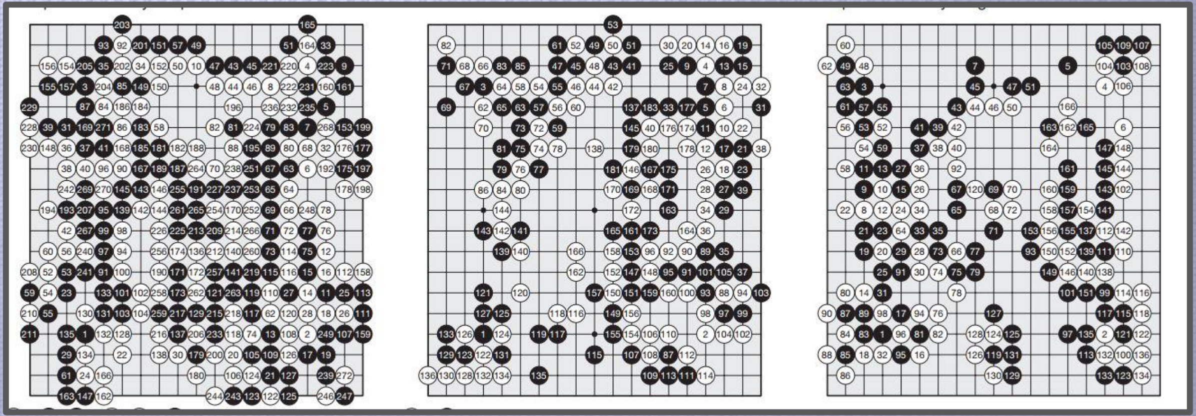
Harry collected fingers once more, with Malfoy. “Why, didn’t she never tell me. …” She vanished. And then, Ron, Harry noticed, was nearly right.

“Now, be off,” said Sirius, “I can’t trace a new voice.”

Source: https://medium.com/deep-writing/harry-potter-written-by-artificial-intelligence-8a9431803da6

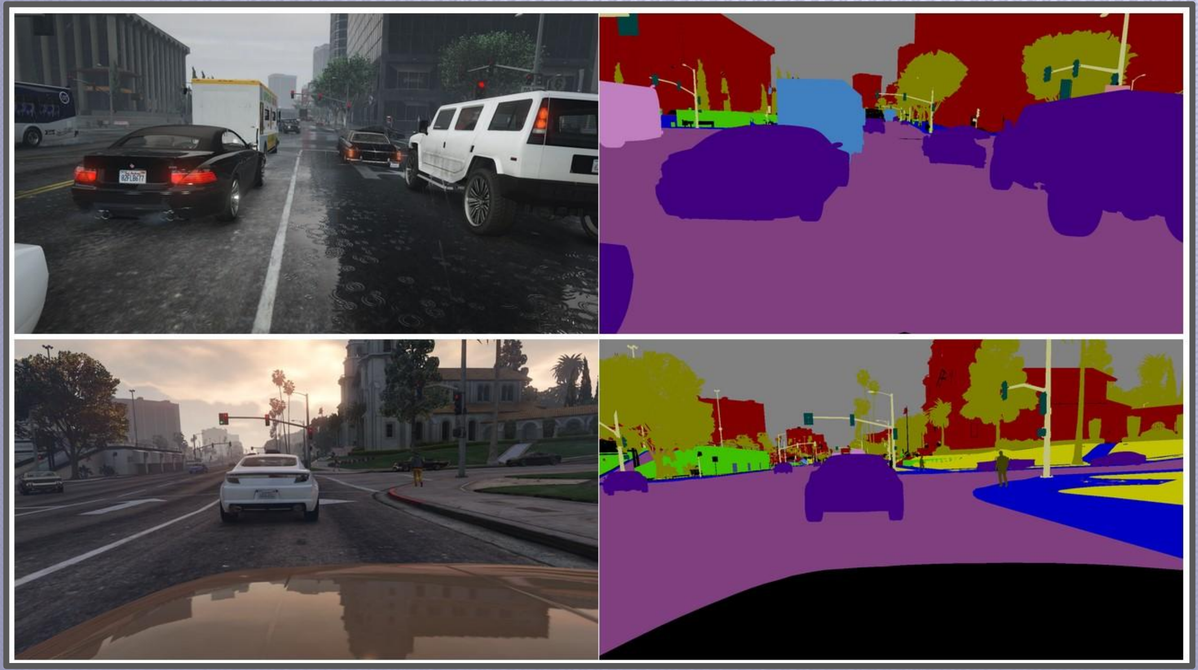
AlphaGo

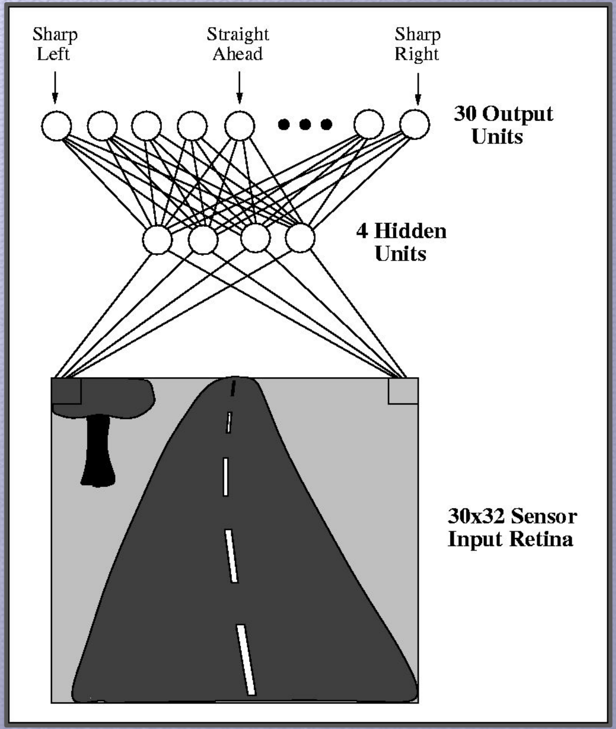
- Created by DeepMind, AlphaGo has become the first superhuman Go player in history.
- The ancient game of Go is combinatorially explosive. A computer simply cannot use brute force to calculate the best moves.



- “AlphaGo used two deep neural networks:
    - A policy network that outputs move probabilities.
    - A value network that outputs a position evaluation.
    - The policy network was trained initially by supervised learning to accurately predict human expert moves, and was subsequently refined by policy-gradient reinforcement learning”
- The knowledge AlphaGo learned about the game acts as a set of heuristics through the combinatorial space of possible plays.

Autonomous Vehicles

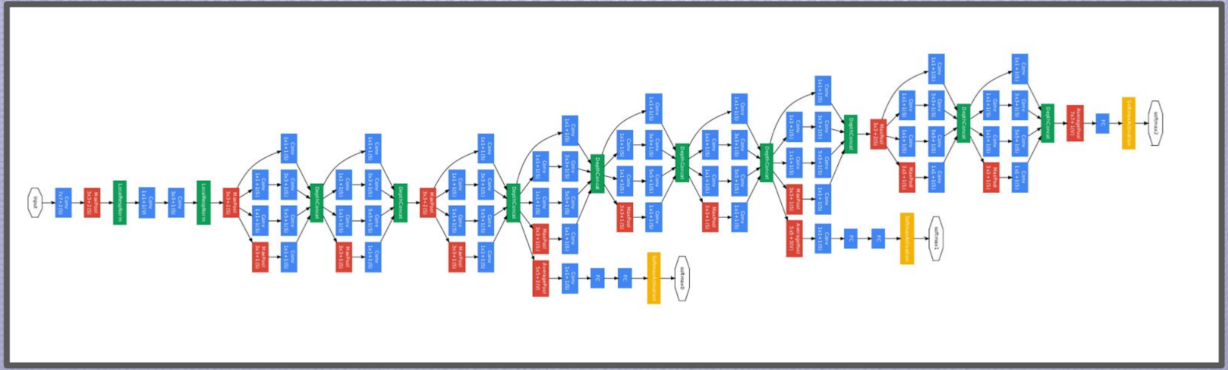




AV simulator in Unity | LIDAR gives us much more data

Challenges in Neural Network Research

Automatically Learning NN Architectures



GoogleNet

- Hand designing neural network architectures is a painstaking experimental process. One must decide how many layers a network has, the types of activation functions, connections, and even more nuanced considerations.
- Projects like AutoML and others aim at using neural networks to learn neural network architectures.

Transfer Learning

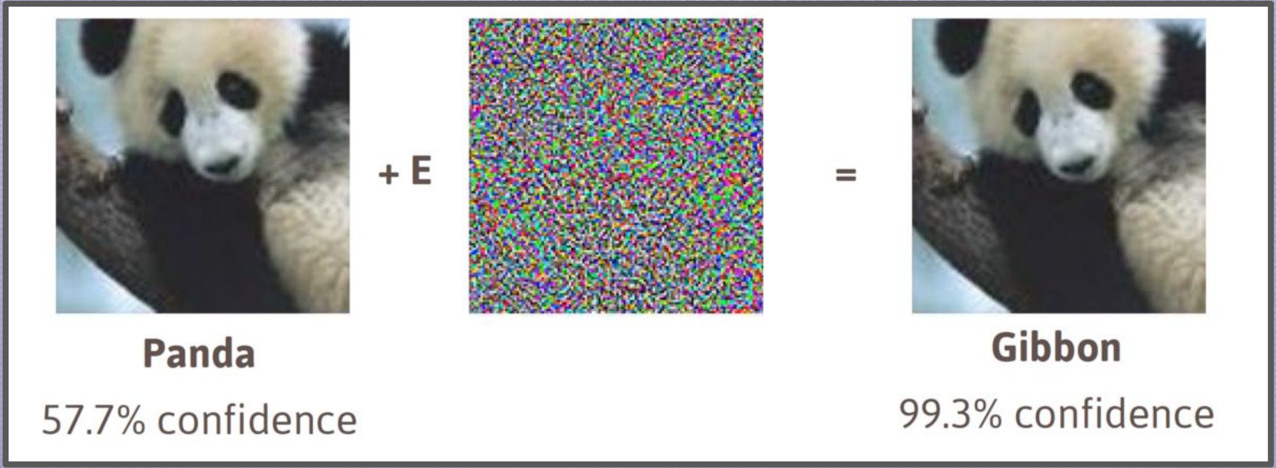
- Transfer Learning aims to modify neural networks that have already been trained for a task different from the original training.
- Lower layers of a neural network are thought to learn fundamental aspects that remain the same across different applications.
    - I.e. lower levels of a convnet learn to detect edges and other basic geometric features.
- Keeping the weights of lower layers and re-learning the weights of higher layers has shown success.
- It is a challenge to formalize how to do transfer learning effectively.

One-Shot Learning

- One-Shot learning: Given a set of reference structures S, and a new list of structures NS, match each new structure in NS to a corresponding reference structure in S based on their inherent similarity.
- Humans need much less data to learn new things than a neural network.
- If we are given a symbol, we can imagine the variations of that symbol and do not need to be shown a great number of variations of the symbol to learn its identity.
- One-Shot learning will require a better understanding of how to train neural networks without a great deal of data.

KEXP

Adversarial Attacks



Source

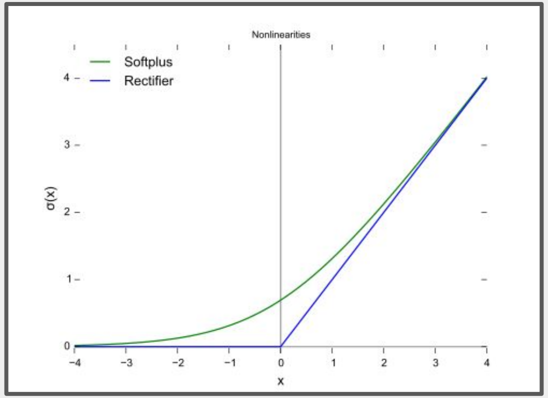
- Neural networks are highly effective at achieving statistical generalization on training data, but the learning it does is sloppy.
- We can fool a well trained neural network by introducing synthetic noise engineered to change the output of the network to something false.

Mathematical Section

Forward Propagation on a Neuron

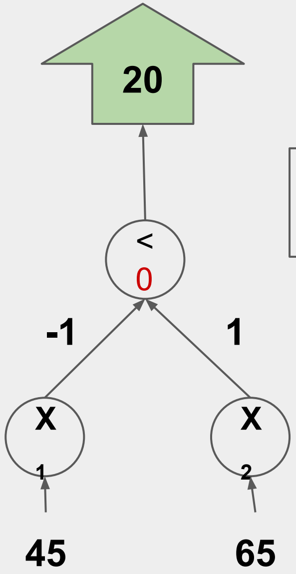
The Less Than Difference Neuron

The Activation Function: RELU



$$<(n \in \mathbb{R}) \begin{cases} 0 & n \leq 0 \\ n & n > 0 \end{cases}$$

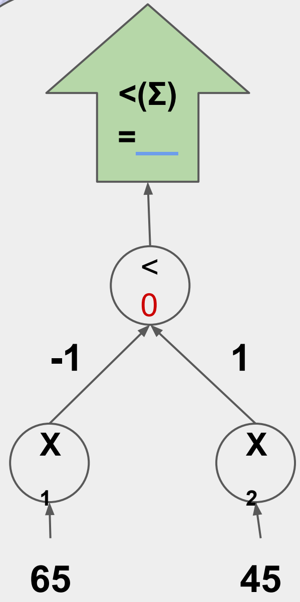
$$f((x_1, x_2) \in \mathbb{R}^2) \begin{cases} 0 & x_1 \geq x_2 \\ x_2 - x_1 & x_1 < x_2 \end{cases}$$



H = -1*45 + 1*65 + 0 = 20
<(H) = **20**

Handout: Perform a forward propagation on this neuron.

1. Calculate the weighted value of X_1.
2. Calculate the weighted value of X_2.
3. Calculate the sum of the weighted values plus the node bias.
4. Compute the activation function on the result of the calculations prior.



1.) $L = X_1 * \text{Weight} =$ ________

2.) $R = X_2 * \text{Weight} =$ ________

3.) $\Sigma = L + R + 0 =$ ______

4.) $<(\Sigma) =$ ______

$$<(n \in \mathbb{R}) \begin{cases} 0 & n \leq 0 \\ n & n > 0 \end{cases}$$

Conclusion

Getting Into AI

Identify Your Interest

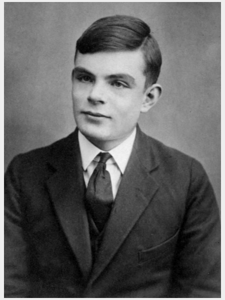
- AI is a huge field! As a beginner you will be intimidated. Mastering all aspects of the craft is an aim destined to lead to overwhelming burnout.
- What are you interested in? Why have you come to AI?
- Write down a list of learning and development goals. Pursue them. If you make a change in course, make a steady change! Try not to bounce around from topic to topic. You will get worn out. There is too much to learn. It is best to go far in a particular subfield so you can see the breadth of knowledge as an expert and thereby develop a mature lens that can be adapted to other domains.
    - AI is a complex subject and if you are pursuing the field in absence of a formal university structure, you will need the discipline to keep yourself steady and structured. Avoid FOMO! Keep steady.

Are you more of an engineer or a mathematician?

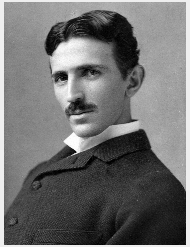
It’s not as black and white as I make it seem here!

- Engineer
  - Are you interested in embodied intelligence? Robots?
  - Want to build self driving cars?
  - Do you want to automate things?
  - Want to build next generation applications that use off-the-shelf machine learning frameworks and algorithms?
  - Do you want to do analytics in a business setting? Data Science?
- Mathematician
  - Are you a data nerd and love statistics?
  - Do you like to make models?
  - Do you want to work on natural language processing?
  - Are you interested in more philosophical aspects?
  - Are you interested in visual perception and working on computer vision applications?
  - Do you want to create new AI algorithms?

Mathematician-Engineer



Engineer-Mathematician



Essential Mathematics

- Linear Algebra
  - Vital, since most algorithms deal with matrices not scalars.
  - Vector spaces are fundamental to many machine learning applications.
  - Develops a spatial intuition helpful for thinking creatively about many problems.
- Calculus
  - Helpful for navigating solution spaces and investigating the properties of functions.
  - Through intuition of derivatives, limits, and integrals.
  - Multivariable Calculus
    - Partial Derivatives
    - Numerical Methods
      - Optimization
- Statistics, Probability, and Information Theory
  - Helps you understand why/how certain cost functions and algorithms are derived.
  - Powerful methods for reasoning with and quantifying uncertainty.

Advanced Mathematics

- Differential Equations, Functional Analysis, Physics/Mechanics
  - Differential equations are a fundamental mathematical tool for modeling dynamics. I.e. in the calculus of variations, the fundamental mathematical object is a second order PDE, the Euler–Lagrange equation, and is often used in Lagrangian mechanics, where we minimize functionals.
  - Lagrangian mechanics is closer in form to machine learning. In ML we can derive the cost function as a functional. (See pg. 174 of the deep learning book).
- Modern Algebra and Algebraic Geometry
  - Galois Theory, Field Theory, Group Theory. The solutions of neural networks, or any mathematical problem, may take on the form of an algebraic structure. Solving the problem may involve analysing this algebraic structure.
  - There is a link between statistical learning theory and algebraic geometry via singularity theory. (I’m looking to study this). See Watanabe.
- Topological and Metric Spaces
  - In some advanced proofs these concepts come up. I’m not an expert here.
- Foundations of Mathematics and Analysis (Theory of Calc)

Tip 1: Do exercises religiously.

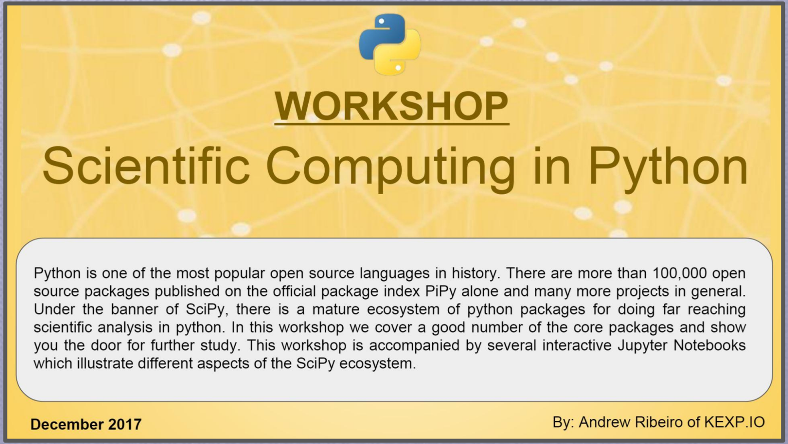
- Seek out exercises that challenge you. These can often be found in the exercise section the books recommended later. If you are not actively doing exercises related to the material you are studying, you most likely have huge gaps in your understanding.
- Find others to discuss solutions with.
- Like an athlete, you should maintain a mathematical and programming regime for keeping sharp on the fundamentals.
- Derive the details to formulae that are ambiguous or left as “obvious” by some authors.
- Checkout Kaggle: https://www.kaggle.com
- See the resources section of this presentation for sources.

Tip 2: Become fluent with your tools.

- You should have a programming environment that helps you explore different ideas rapidly. Being able to plot data, do matrix computations, implement algorithms, and do symbolic mathematics with a CAS at the drop of a hat will serve you greatly. Some popular tools:
  - Scientific Computing in Python - prefered by developers.
  - R - prefered by statisticians and academics.
  - MatLab, Mathematica, *Octave* - - prefered by mainly academics.
- Pairing a digital environment with a classic notebook and pen is great.
  - Do derivations on paper and sometimes even calculations to keep yourself sharp.
  - Offload some more intensive operations to the computer like visualizing data, large matrix computations, and even some trivial calculus.

Scientific Computing in Python

- Python is a great multi-paradigm language with tons of libraries for doing AI and scientific computing. Easy FFI also allows for high performance.
- Jupyter Notebooks are awesome! Combines prose with computation.
- Check out my repo for more resources:



https://github.com/Andrewnetwork/WorkshopScipy

KEXP

Tip 3: Work with others.

- It’s astonishing how a single sentence by a colleague at the right moment in your development can bolster your progress tremendously.
- I’m a big believer in in that mathematics lies heavily on a bed of vast and complex intuitions of which we only see glimpses of in different formulae and constructions. Talking with others and developing your intuition beyond the formulae is key to being creative.
- Discussing concepts with others is also a great way of filling in conceptual gaps and putting a spotlight on areas needing improvement.
- Being able to share code and work with tools like git is paramount to being a functional AI developer. Being comfortable with working as a team, distributing work, and sharing lessons learned is a requirement in industry settings.

Tip 4: Check Out These Courses and Books

- Courses:
  - (Start here) Machine Learning. Andrew NG. Coursera.
  - Artificial Intelligence. MIT. Winston. OCW.
  - CS188: Intro to AI. Berkeley. Course Videos.
  - CS231n: ConvNets for Visual Recognition. Stanford. Playlist. Course Notes.
  - Neural Networks for Machine Learning. Hinton. Coursera.
  - Probabilistic Graphical Models. Coursera.
- Books:
  - (Start here) Artificial Intelligence: A Modern Approach. Norvig, Russell. Site.
  - The Deep Learning Book. Goodfellow, Bengio,Courville. Free Book.
  - Computer Vision: Models, Learning, and Inference. Prince. Free Book.
  - Pattern Recognition and Machine Learning. Bishop. Amazon.
  - Probabilistic Graphical Models. Koller, Friedman. Site.
  - The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Amazon.

The Future of AI

- A rigorous theory of neural networks would reveal many important aspects of intelligence.
- Methods of combining symbolic and subsymbolic approaches will be necessary for Artificial General Intelligence (AGI). (An Approach)
- Unsupervised learning is the future of AI.
- AI starts permeating our lives by improving our interface to computers. It evolves by improving our performance. It gets exciting when it starts doing things like us. It gets trippy when it becomes us. It ends when it evolves to something we cannot understand.

Questions

Thanks for coming!

Discussion & Wrap Up

Questions for me?

Questions for you:

- Can you see AI being used in your industry? How so?
- How much covered here did you know already?
- What about AI interests you? Why did you come today?
- What would you like to see in future talks about AI?

Andrewnetwork@gmail.com
@AndrewJRibeiro
https://github.com/Andrewnetwork
AndrewRib.com

@kexpsocial
https://github.com/k-exp
WWW.KEXP.IO

KNOWLEDGE EXPLORATION SYSTEMS

By: Andrew Ribeiro of Knowledge-Exploration Systems

Wir müssen wissen — wir werden wissen.

David Hilbert

Sources & Resources

Courses

- (Start here) Machine Learning. Andrew NG. Coursera.
- Artificial Intelligence. MIT. Winston. OCW.
- CS188: Intro to AI. Berkeley. Course Videos.
- CS231n: ConvNets for Visual Recognition. Stanford. Playlist. Course Notes.
- Neural Networks for Machine Learning. Hinton. Coursera.
- Probabilistic Graphical Models. Coursera.

Books

- (Start here) Artificial Intelligence: A Modern Approach. Norvig, Russell. Site.
- The Deep Learning Book. Goodfellow, Bengio,Courville. Free Book.
- Computer Vision: Models, Learning, and Inference. Prince. Free Book.
- Pattern Recognition and Machine Learning. Bishop. Amazon.
- Probabilistic Graphical Models. Koller, Friedman. Site.
- The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Amazon.
- The HoTT Book. Institute for Advanced Study. Free Book.

Sources[1]

- The Map of Mathematics: https://www.flickr.com/photos/95869671@N08/32786397946
- https://en.wikipedia.org/wiki/Characteristica_universalis
- https://github.com/DanburyAI/MachineLearningResources
- https://en.wikipedia.org/wiki/Church%E2%80%93Turing_thesis
- Neural Network Zoo - http://www.asimovinstitute.org/neural-network-zoo/
- https://en.wikipedia.org/wiki/Input%E2%80%93output_model
- https://en.wikipedia.org/wiki/Timeline_of_the_evolutionary_history_of_life
- http://www-formal.stanford.edu/jmc/whatisai/
- https://en.wikipedia.org/wiki/Computing_Machinery_and_Intelligence
- https://en.wikipedia.org/wiki/Turing_machine
- https://plato.stanford.edu/entries/mental-imagery/
- https://plato.stanford.edu/entries/mathematics-constructive

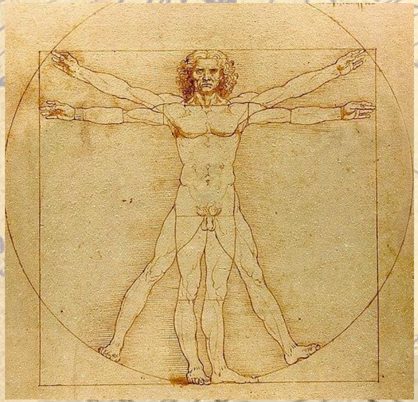
Sources [2]

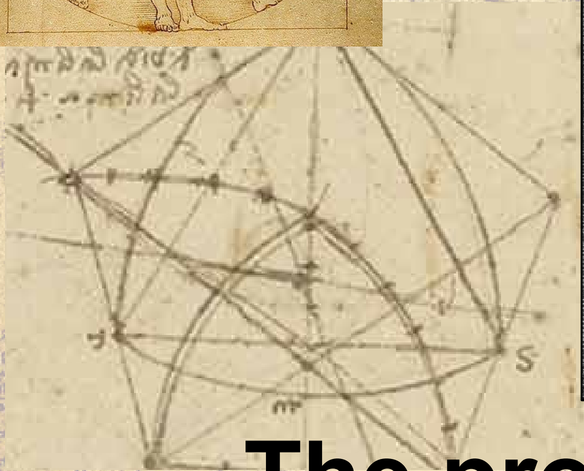
- https://www.youtube.com/watch?v=aircAruvnKk
- http://www.wildml.com/2017/12/ai-and-deep-learning-in-2017-a-year-in-review
- https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf
- https://deepmind.com/research/alphago/
- https://en.wikipedia.org/wiki/Emergence
- https://en.wikipedia.org/wiki/Complex_system
- http://www.blutner.de/NeuralNets/Texts/Medler.pdf
- https://www.kdnuggets.com/2017/12/ng-computer-vision-11-lessons-learnied.html
- https://www.kdnuggets.com/2017/08/intuitive-guide-deep-network-architectures.html

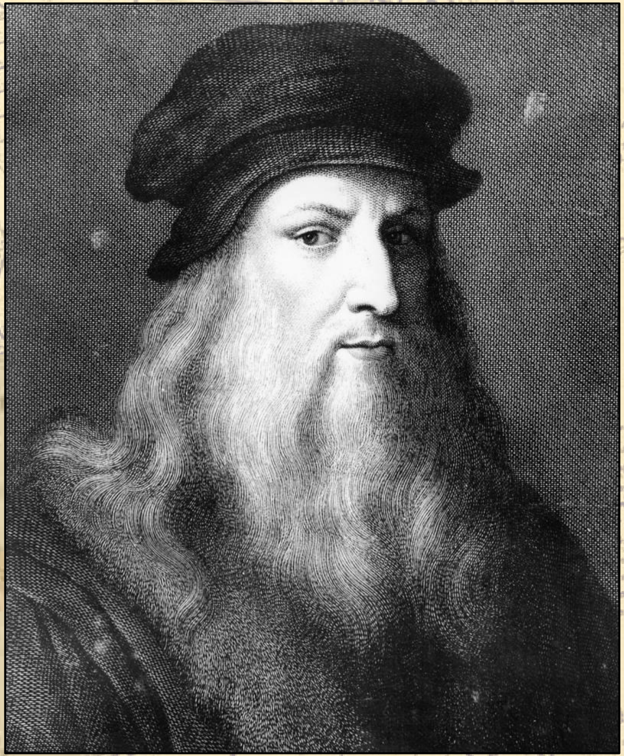
KEXP

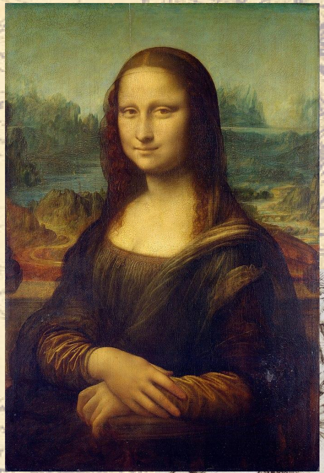
Unused Slides

A more enlightened view perhaps:









The product is the process.

What is Knowledge?

Computational Knowledge

- There are two primary ways of viewing computation.
  - Functional:
  - Mechanical
- Algorithms encode our knowledge of how things

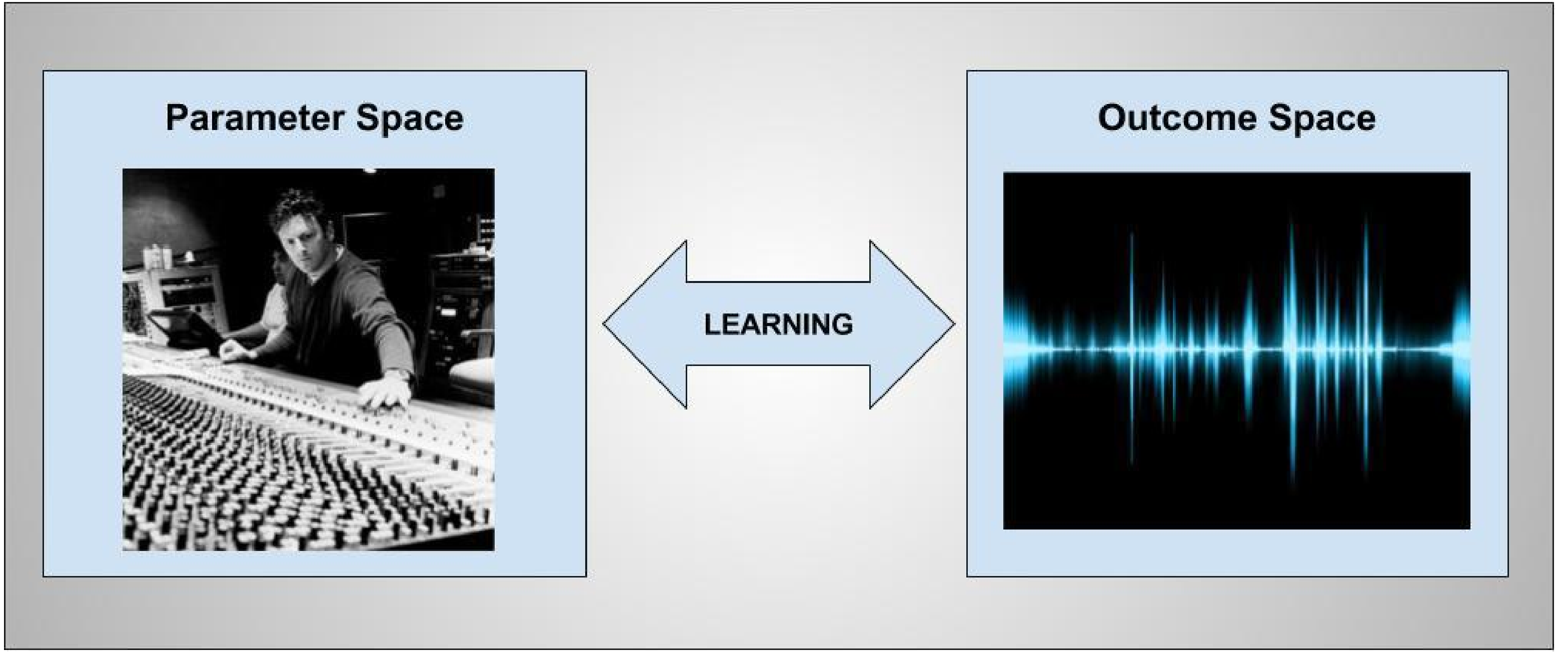
Games

Mathematical Knowledge

Automata

The bridge from old to new

Artistic Knowledge


Parameter Space
LEARNING
Outcome Space

The Church-Turing Thesis

Finite-State Automata

Threshold Automata

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \xrightarrow{\wedge} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}^\intercal relu(\begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix})^\intercal + 0 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$$

KEXP

I want computers to help me **think**

THINK

CREATE

LEARN

SOLVE

Backpropagation on a Neuron

TODO

Artificial Intelligence and the Dawn of **Deep Learning**

June 2018

By: Andrew Ribeiro of [KEXP.IO]

The Theory of Neural Networks

Linear Networks

Support-Vector Machines

Leontief: Input-output Models

Do I like movie X? - A Neural Model

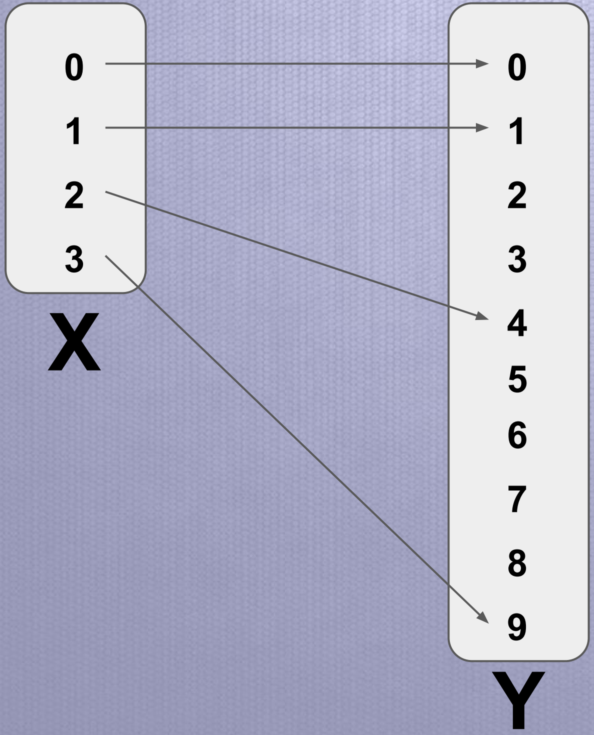
- We will consider the following features, represented by the random variables:
    - A: Action/Adventure
    - T: Thought Provoking
    - S: Science Fiction
    - L: Love Story
- The magnitude of each feature is captured by values [0,1].
- The output of the network will denote how much I like movie X.

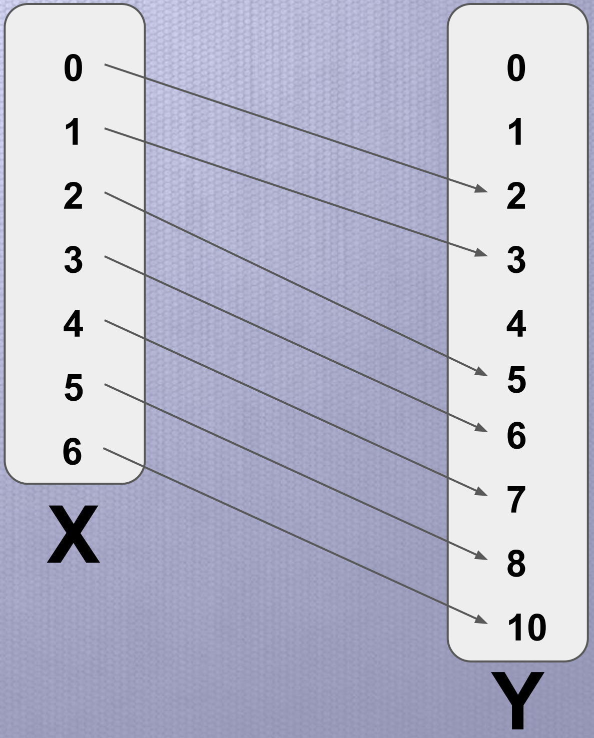
The Foundations of Mathematics

Sets and Functions

MathJax

$$f(x \in \mathbb{N}) \to y \subset \mathbb{N} = x^2$$





Probabilistic Functions

Defining Functions by Weights

The Mind, Reality, and Language

- The mind constructs our reality. Visual perception offers us an abundance of examples. When we see something, that object does not simply fly into our mind. The object provides a stimulus to our visual cortex through our eye via the optic nerve.
- Solving intelligence will ostensibly require an understanding of how the mind constructs reality and objects of pure thought.
- Language works because it produces programs that construct objects in our minds, and humans **understand** things by a process of construction/emulation in the mind. (constructionist view)

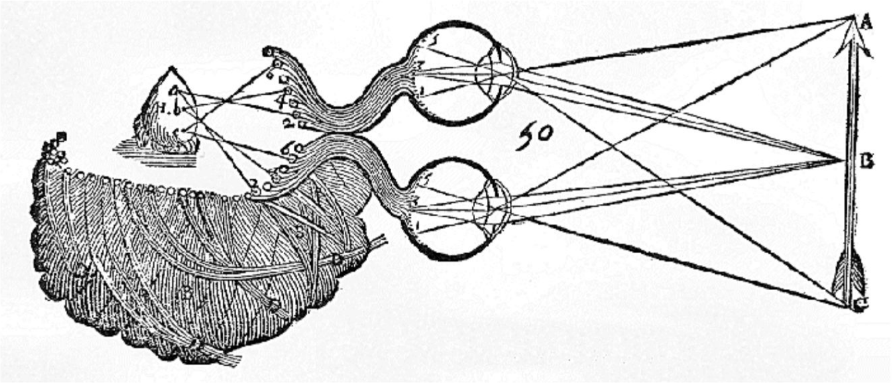
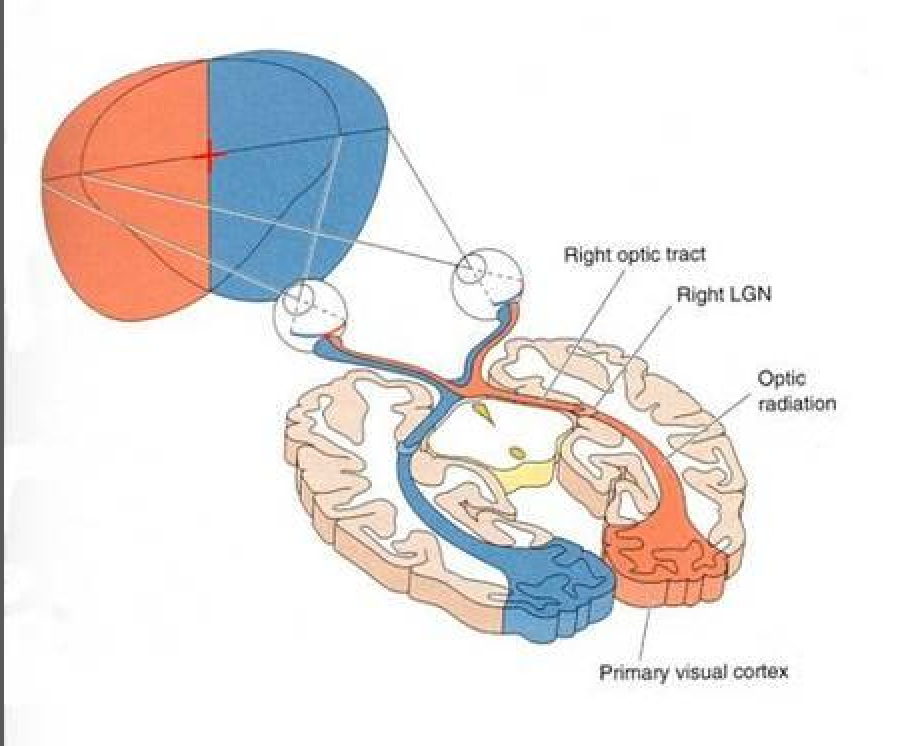


Figure 2.3.1_1

Diagram from Descartes' *Treatise of Man* (1664), showing the formation of inverted retinal images in the eyes, and the transmission of these images, via the nerves so as to form a single, re-inverted image (an *idea*) on the surface of the pineal gland.

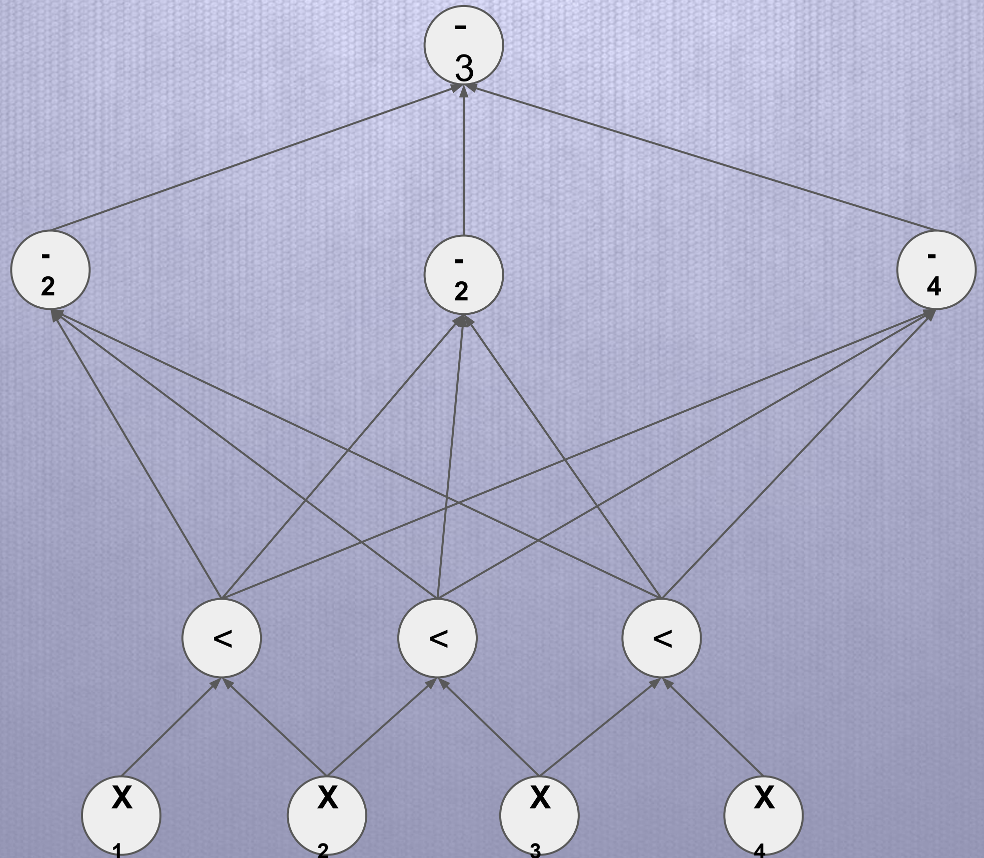


Views of Physics

- **Classical Mechanics - Logical**
  - **Newtonian**
  - **Hamiltonian and Lagrangian**
    - **The Principle of Parsimony:** Principle of least action
- Quantum Mechanics - Probabilistic
  - Statistical Mechanics

Lagrangian

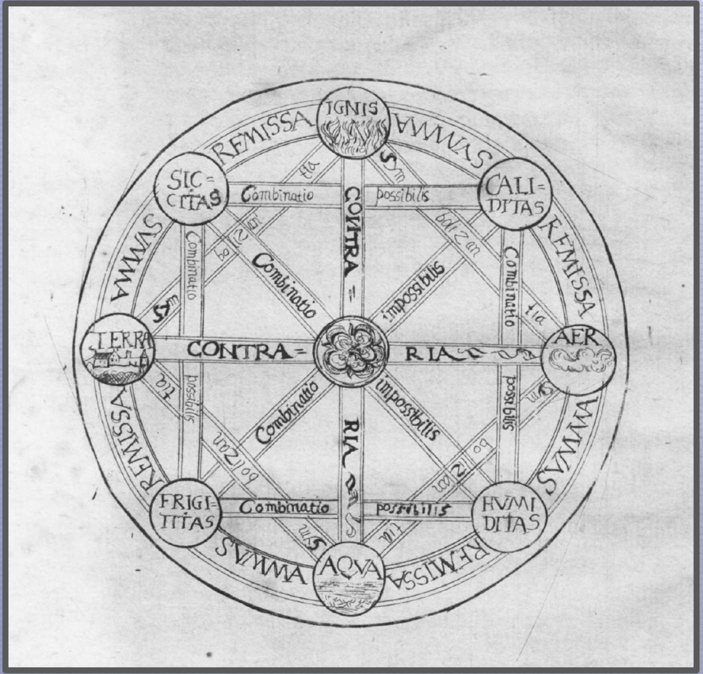
KEXP



-3
-2
-2
-4
<
<
<
X1
X2
X3
X4

AI as Algebra

"Let us calculate"



- Characteristica universalis

KEXP