Artificial Intelligence

What is this stuff, really?

Presentation to Phoenix IEEE Computer Society
11 April, 2018
By Dr Brad Morantz

Note

I am a mathematical decision scientist and therefore my focus is on the problem at hand and how I can best solve it. My focus is NOT the computer itself.

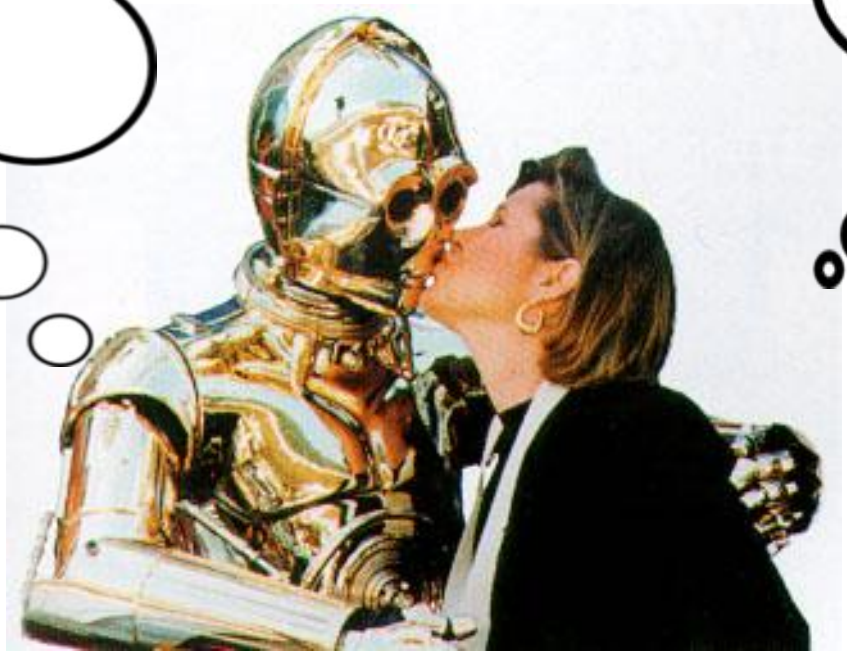
I will discuss what AI is, where it came from, and some examples. This could easily be a one or two semester class, so we will cover what we can in one evening.

It is not important which computer language is used so I will not even touch upon this topic. This is about knowledge, intelligence, and learning, but not programming.

An old saying: If the only tool in your toolbox is a hammer, then every problem looks like a nail.

Star Wars™

If I had any REAL brains would I be doing this?



I hope that I don't short out any of his circuits.

Is this intelligent?

What is Intelligence?

- Who knows what it is? (a big question!)
- Ability to understand or reason (dictionary)
- Mental ability: learning, problem solving, abstract thinking, spatial manipulation, language acquisition, & reasoning (encyclopedia)
- Herb Simon:
  - Involves associations, pattern recognition, inference, experience, & intuition
- 1948 Conference
  - It is what is measured by IQ tests

What does “Artificial” mean?

- Random House College Dictionary:
  - Produced by man
  - Made in imitation or as a substitute
  - Simulated

- Examples
- Artificial Chocolate
  - May look and taste like chocolate, but it is **not**
- Hot dogs

  - Soy dogs look like hot dogs, kind of taste like them, are definitely healthier, but contain no meat. Are NOT really hot dogs!

Then what is Artificial Intelligence?

Combining the terms

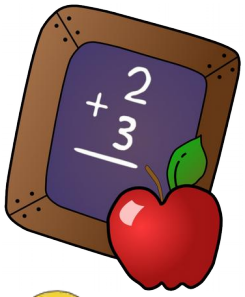
- Simulated ability to understand, reason, and problem solve, or at least appear to (my definition)

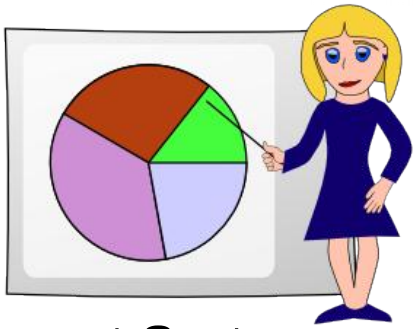
  Mimicking "cognitive" functions that humans associate with other human minds "learning" and "problem solving"- Russel & Norvig
- Ability of a computer to perform tasks (that human intelligence is capable of doing) such as reasoning and learning. [Ed: or appear to] (McGraw-Hill computer Handbook)

Note: if it could really do that, then would it be “artificial”?

Any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals [Poole, Mackworth, & Goebel]

What are our goals?





- Solve problems
- Improve performance
- Increase profits
- Forecasting
- Better decisions
  - DSS Decision Support Systems
- Model biological to further understanding
  - Forgotten but important
  - Test pharmaceuticals
  - Look for cures
  - Understand the brain/CNS

Intel’s AI Applications

- Partner w/ Ferrari for driving experience
  - AI to “better synthesize data”
- Make personal finance easier
- Advance brain research with Princeton Neuroscience Institute
- Analyze whale health to help protect ocean health w/Parley
- Solve business problems
  - Discover fraud
  - Identify manufacturing errors
  - Optimize crop yields

This is just an example of applications

Sample Applications

- Mycin
  - Expert system that helps doctors to diagnose infectious blood diseases
- Teresius
  - Expert system to help with investments
- Microsoft Office™
  - Uses AI to help correct mistakes
  - To do what it thinks is best
- My work in Neural Network time series forecasting
  - forecasting CD rates and financial data

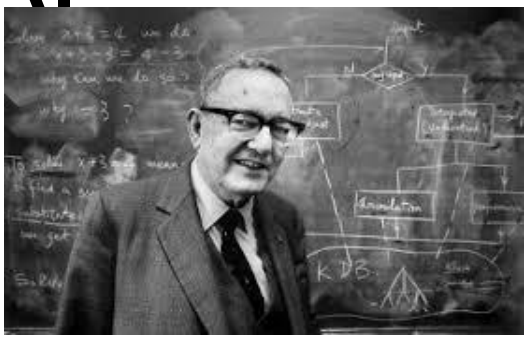
Origin

- 1941 Research in 'machine intelligence' UK
- Alan Turing *“Intelligent Machinery”* - 1947
- Did not define intelligence, instead asked if 'machines can do what we do'.
- Manipulating symbols is not understanding
- Turing test - 1950

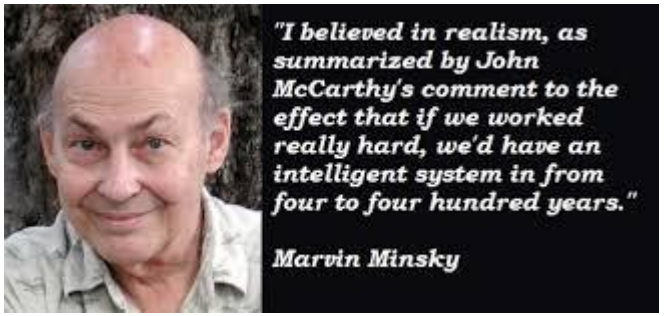
Turing Test

- The test
  - Alan Turing *Computing Machinery and Intelligence,* 1950
  - Started with “Can machines think?”
  - Test he devised:
    - Two people having discussion over teletype (Ed: it was 1950)
    - One of the men replaced by machine
    - Can we detect that it is machine, not man?

Two Fathers of AI



- Herb Simon – CMU (1916 - 2001)
  - Psychology & Computer Science
  - *Human Problem Solving (*with Newell*)* – 1972
- Marvin Minsky – MIT (1927 - 2016)
  - Artificial Intelligence Laboratory
  - *Society of Mind* – 1985
  - *Perceptrons* (with Papert)



- These are just a few of their many accomplishments & papers
- There were also many others, too many to list

Current Computer “AI”

- Artificial Neural Networks
- Genetic Algorithms (evolutionary computing)
- Data Mining/ Machine Learning
- Expert Systems
- Case Based Reasoning
- Synthetic Immune Systems
- Fuzzy Logic
- Hybrid
- Generative Adversarial Networks
- Cognitive architectures (SOAR, Cyc, Opencog, ACT-R, etc)

Artificial Neural Networks

- Computer emulation of biological brain (ANN)
- Massively parallel system
- General data driven function approximator
- Intel Nervana™ Neural Network Processor (NNP)
- Functions performed
    - Pattern recognition
    - Classification
    - Forecasting/linear & nonlinear* regression
    - Brain emulation

*Logistic, Sigmoid, hyperbolic tangent, etc

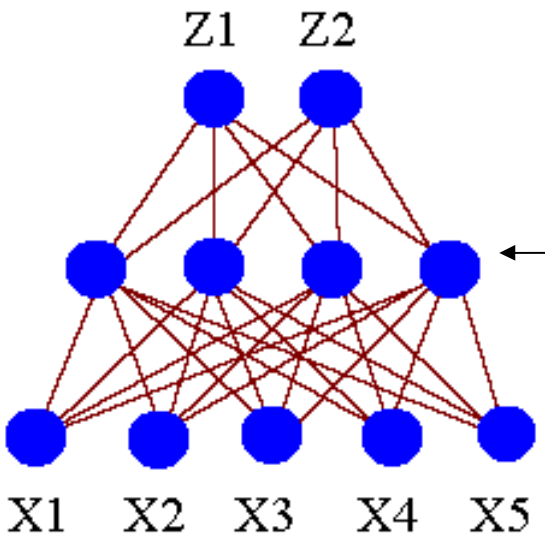
About ANNs

- General function approximator
  - Imitates performance of original
  - Does not duplicate model
  - Does provide near or approximate results
  - It maps input to output
- Data driven
  - Does not understand causal model
  - Learns input to output relationship
  - Learns from supplied training data

It does not think, it just maps

Feed Forward Neural Network

Output



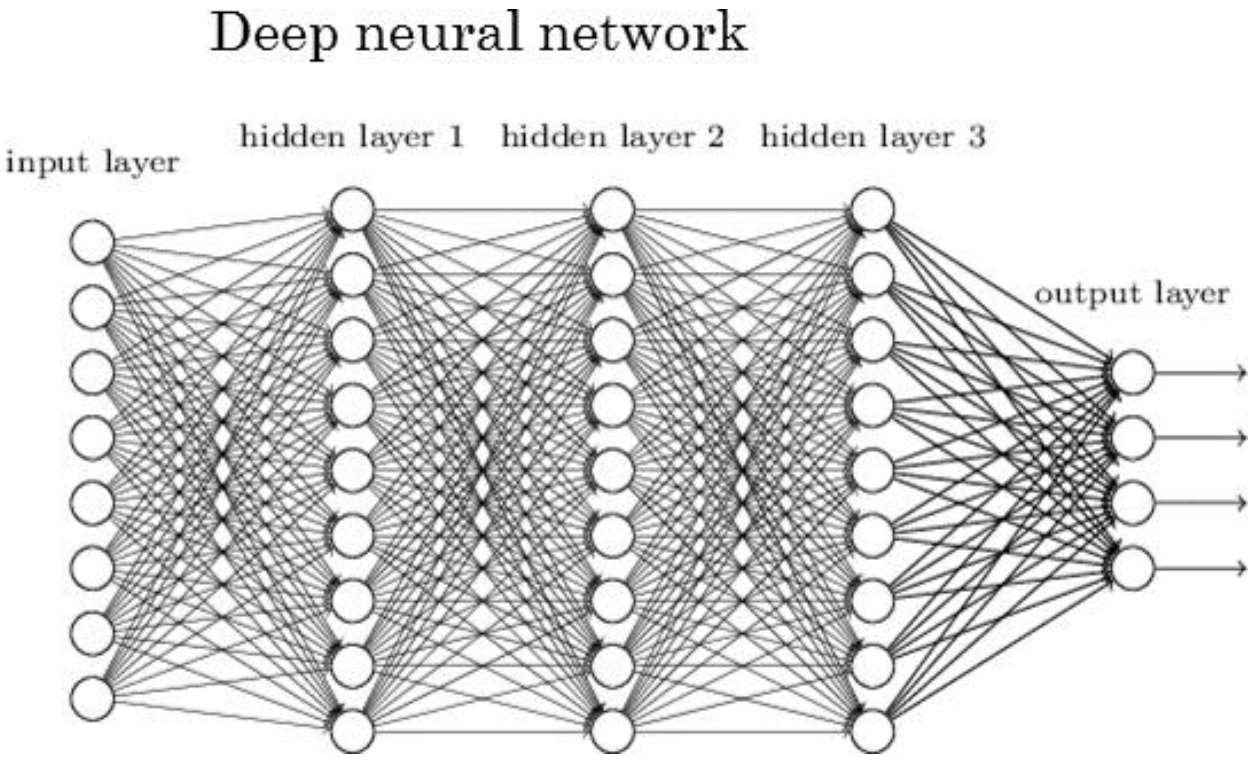
Hidden layer

(Can have more than one)

input

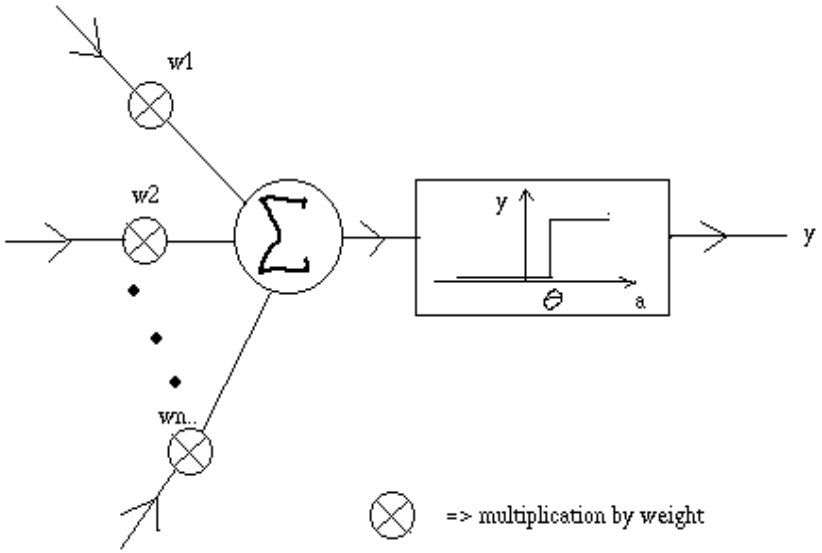
“Deep Learning” has one or more hidden layers
Increases complexity of equation

Deep Learning ANN



Notice the extra hidden layers
Similar to organization in biological brain
Every connecting line is a coefficient
Every node is a variable
Imagine how complex the equation

Model of Individual Neuron

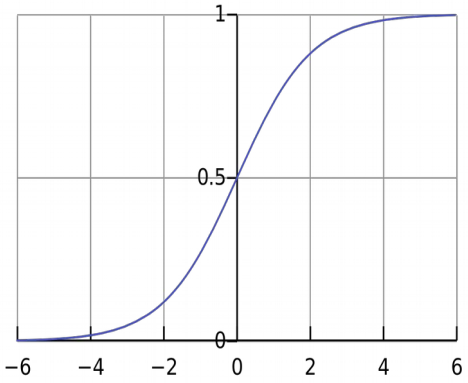


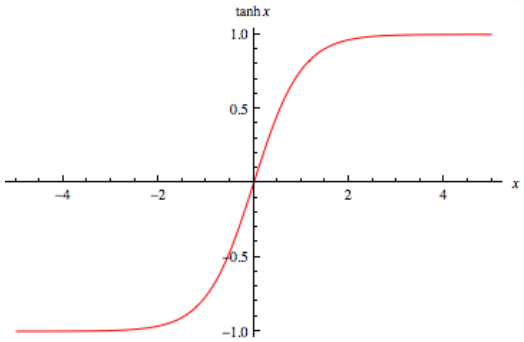
- Input is a large number of weighted outputs from nerves or other neurons
- It sums the weighted inputs
- If the sum is greater than a threshold, then it fires
- In Human, can have as much as 10,000 connections (some say 100,000)

Activation Functions

- Weights can be positive or negative
- Negative weight inhibits neuron firing
- Sum = $W_1N_1 + W_2N_2 + \ldots + W_nN_n$
- If sum is negative, neuron does not fire
- If sum is positive neuron fires
- Fire means an output from neuron
- Can be non-linear function
- Some models include a threshold

Neuron Activation





- Linear
- Sigmoidal
  - $1.0/(1.0+e^{-s})$ where s = Σ inputs
  - 0 or +1 result
- Hyperbolic Tangent
  - $(e^{s} – e^{-s}) / (e^{s} + e^{-s})$ where s = Σ inputs
  - -1 or +1 result
- Also called squashing or clamping function
  - Because it takes a large value and compresses it
  - Adds the non-linearity to the process

Functions of an ANN

- Prediction and Time Series Forecasting
  - Like regression, but not constrained to linear
- Classification
  - Find which class is the closest match
- Pattern Recognition
  - Fined tuned classification – exact match
- Self organizing map for clustering
- Not constrained to linear or Gauss Normal distribution
- Also used for modeling biological neural network in medical research

Advantages of an ANN

- No Expert needed
- No Knowledge Engineer needed
- Does not have bias of expert
- Can interpolate for all cases
- Learns from data/facts
- Can resolve conflicts
- Variables can be correlated (multicollinearity)

More Advantages of ANN

- Learns relationships
- Can make good model with noisy or incomplete data
- Can handle non-linear or discontinuous data
- Can Handle data of unknown or undefined distribution
- Data Driven

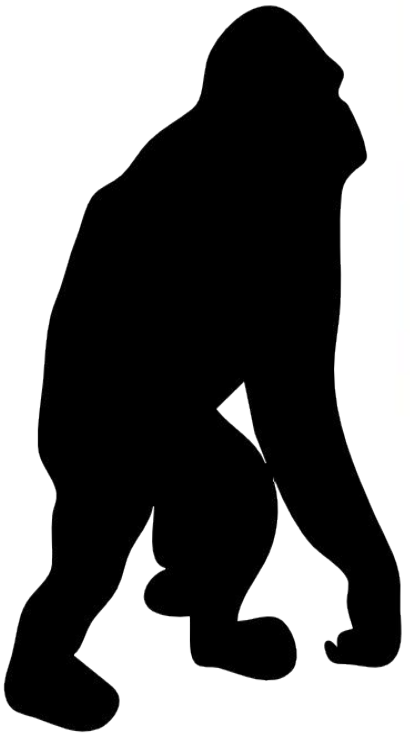
Disadvantages of ANN

- Black Box
  - don’t know why or how
  - not sure of what it is looking at
- Operator dependent (improvements ongoing)
- Do not have knowledge in hand
- Just a mapping of input to output, no modeling of system

** Many of these disadvantages are being overcome*

Genetic Algorithms

- John Holland and Schema Theorem, 1975
  - Organisms of greater fitness appear with exponentially greater frequency as chromosomes are replaced
- Imitates natural evolution
  - Also called evolutionary computing
  - Modeled on natural selection
  - Survival of the fittest
- Exploited search in hyperspace (N space)
- Near optimal solution for complex problems



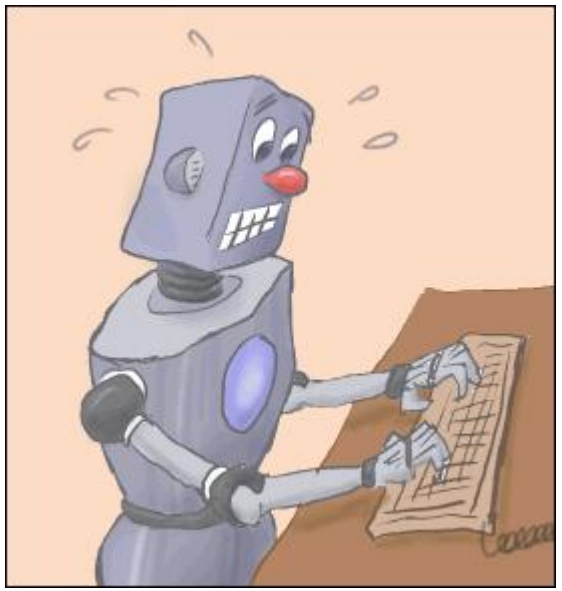
General Operation

- Problem is defined
- Each potential solution is a chromosome
- Generate population of chromosomes
- Keep and mate the most fit
- Let the least fit die off (no bailouts or too large to fail)
- Each successive generation gets stronger

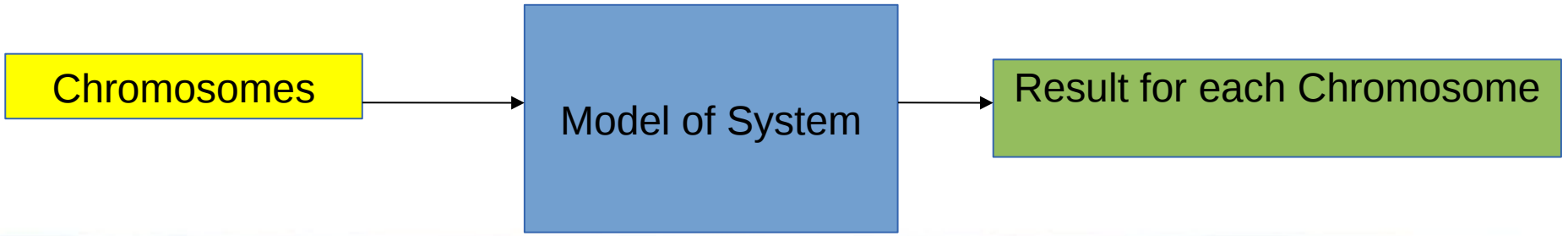
Example chromosome

0 1 1 0 1 0 0 1 1 1 0 1 0 1 0 1 1 0

Fitness Function

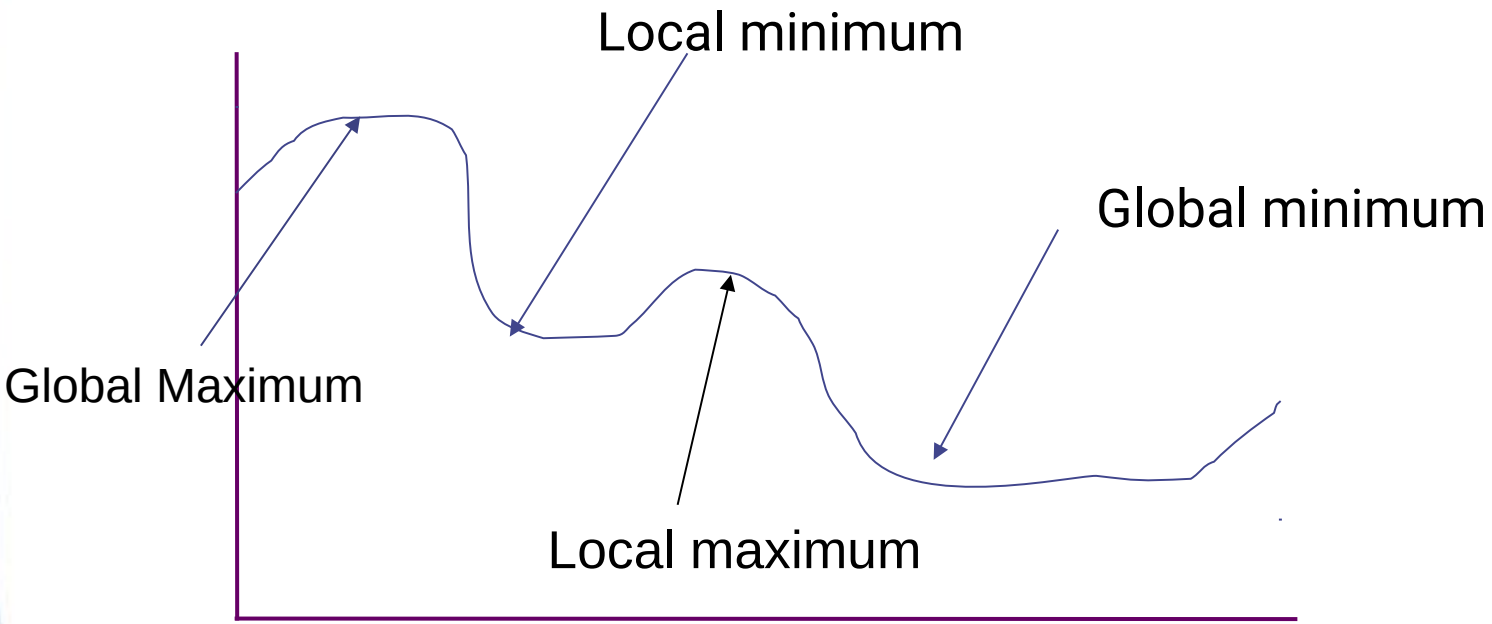


- Evaluates each chromosome
- Evaluates system performance with solution prescribed by organism
- Calculates figure of merit for each
  - Assigns a metric to performance in area of interest
- Ordering values will show which are best performers or most robust (and which the worst)



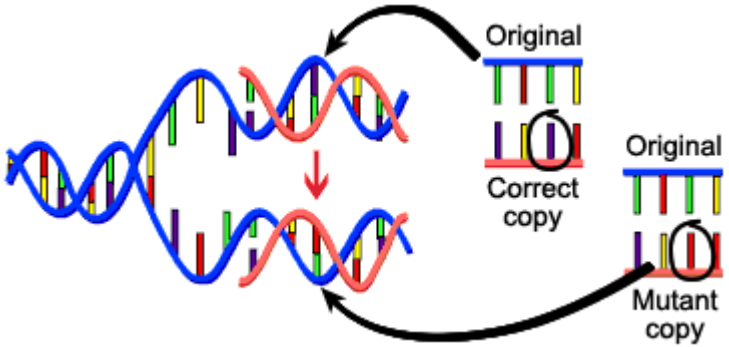
Local Minima

- Major problem: can get trapped in plateau when need to find global minima



Mutation

- An allele just randomly changes
- Expands the search space
- Introduces new features
- Random process
- Accomplished by randomly changing a gene
- Empirically shown that optimum rate is 1.5% to 3%



Pros and Cons

- Large overhead
- Handles big problems
- Can handle non-linear
- Also can be used to find worst solution
- Fitness function may take hard work
- The real intelligence is in the making of the fitness function

A First

John Koza

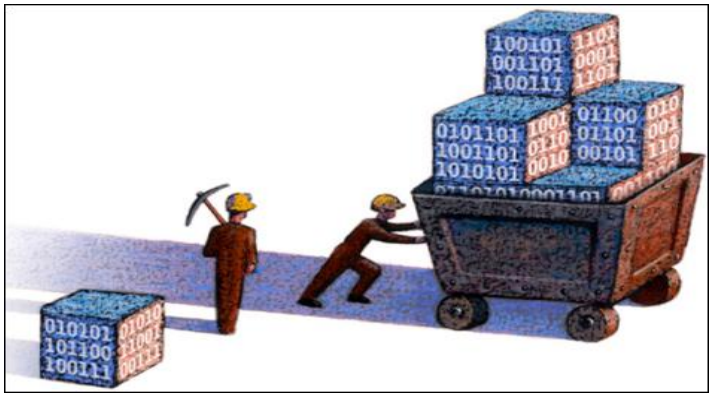
- Studied under John Holland
- Made invention machine
- 1000 networked computers
- For optimizing many things
- Machine created program for optimizing factories
- One of first patents for IP (intellectual property) for a non-human
- Virtually no human guidance

Data Mining Definition

- Is the nontrivial extraction of implicit previously unknown and potentially useful information and data. (James Buckley)
- Machine Learning employs search heuristics to uncover interesting and systematic relationships in data. (Dhar & Stein)
  - Going into a data file and looking for patterns and relationships
  - Use of historical data to find patterns and improve future decisions
  - Also called KDD
    - Knowledge discovery from databases

Data Mining

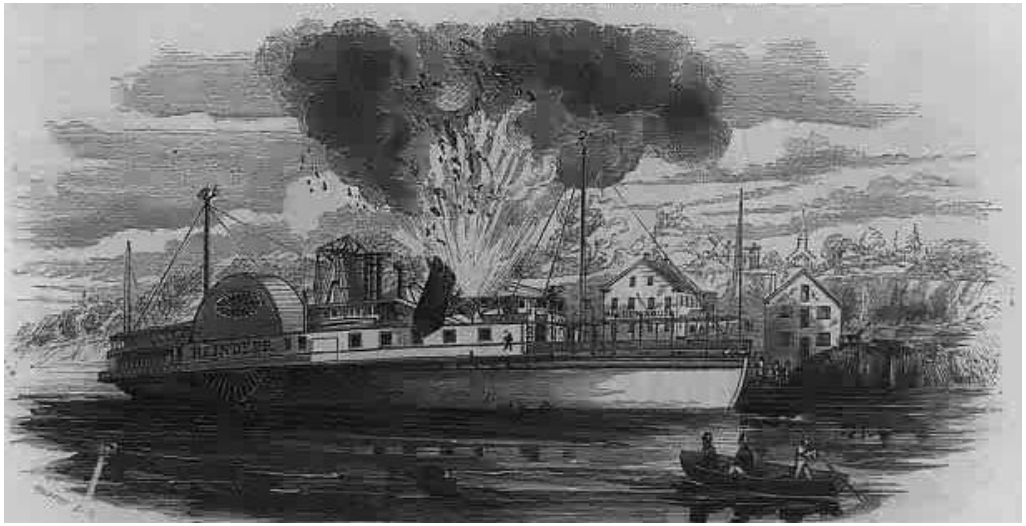
- Also known as Machine Learning
- Tons of data available today
- Look into the data
  - No preconceived ideas
  - Look and see what you find
  - Look for patterns



- Today, people search data for specific things
- Heavily operator dependent
- Try statistics first, then SOM or cluster analysis
  - SOM is Self Organizing Map (a type of neural network)
  - Cluster analysis groups things by proximity in variable space
  - Methods to group observations upon dimensions
- Where is the intelligence?

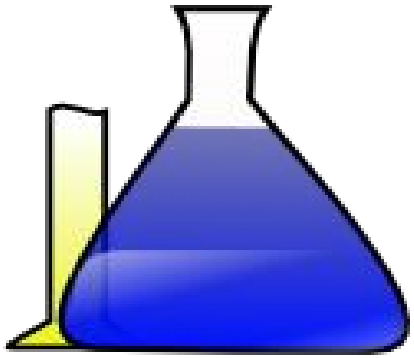
Original

- US Navy
- Boilers on boats exploding
- Examined data
  - Found pattern
  - Prevented future boiler damage



Applications

- Marketing
  - Target marketing
  - Store layout
  - Better understanding
- Medical
- Criminal
- Financial
- Science
- Etc.



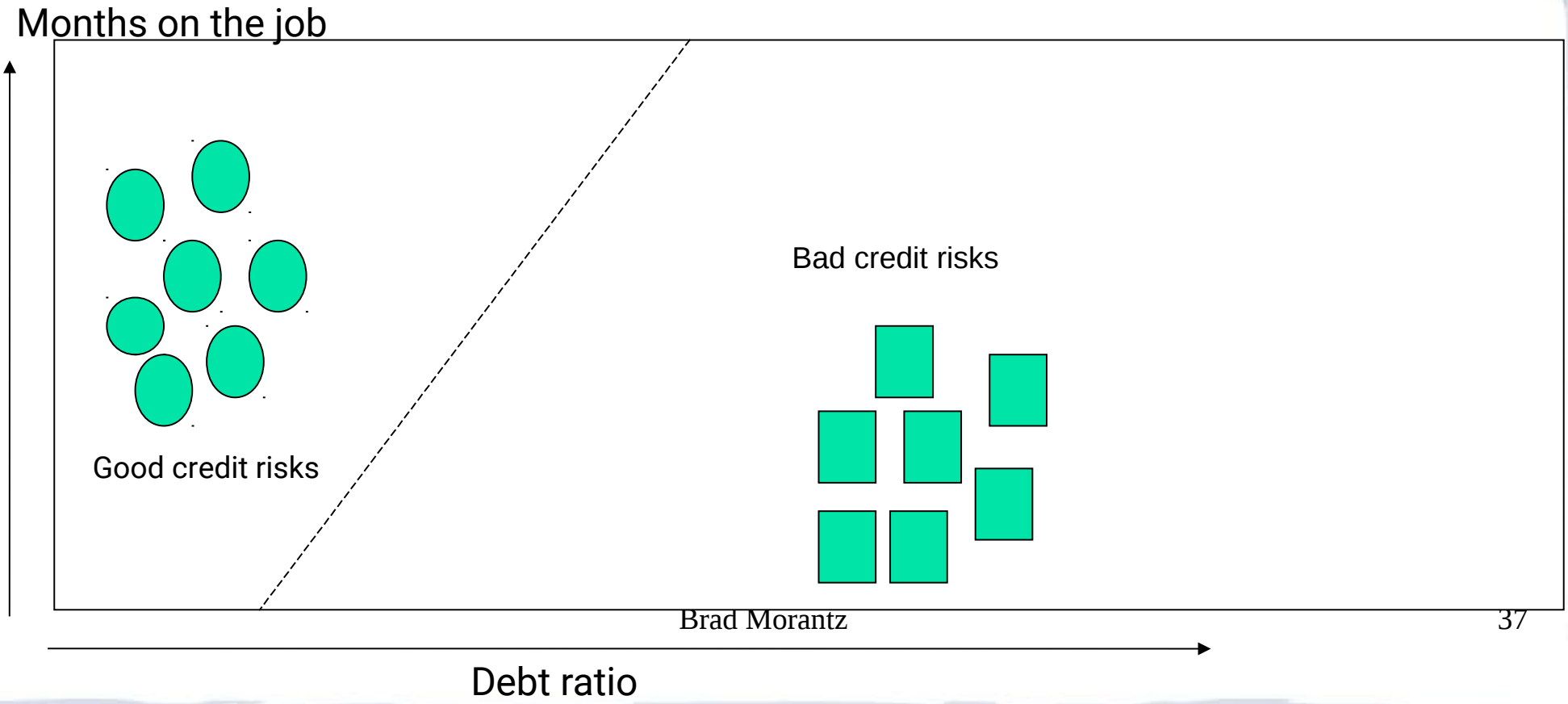
Methods

- Rule Induction
- Similarity Engine
- OLS & logistic regression
- Data visualization
- Bayesian Classification & networks
- Clustering
- Specialized programs (Big $$)
  - Clementine
  - Enterprise Miner

Plot

Plot data and put in breaks to split it

Support Vector Machine (SVM) also does this
- Width of fence indicates separation



Expert Systems

- Just like having a subject expert
- Creates a Decision Tree
- Stored in a set of “If.. then..” rules
- Consists of
  - Rule base
  - Inference engine/rule interpreter
- Get rules from Human Expert
- Knowledge engineer converts knowledge into rules
- System seems intelligent when using it
  - It is the intelligence put into it
- Example
  - If this is a corner, then must go into second gear
  - My chocolate chip cookie program

Using an Expert System

- Steps
  - Hire an expert
  - Hire a knowledge engineer
  - Create rule set
  - Apply problem
- Limitations
  - Can only answer problems that it has already seen
  - Contains biases of expert
  - Where is the intelligence?

Advantages

- Knowledge Based
- Works without expert present
- Consistent performance
- Can explain why (very major point)
- Rules can easily be added
- Rules can easily be modified

Disadvantages

- Does not learn
- Requires an expert
- Requires a knowledge engineer
- Must have rules for all possibilities
- Contains bias of expert

Case Based Reasoning

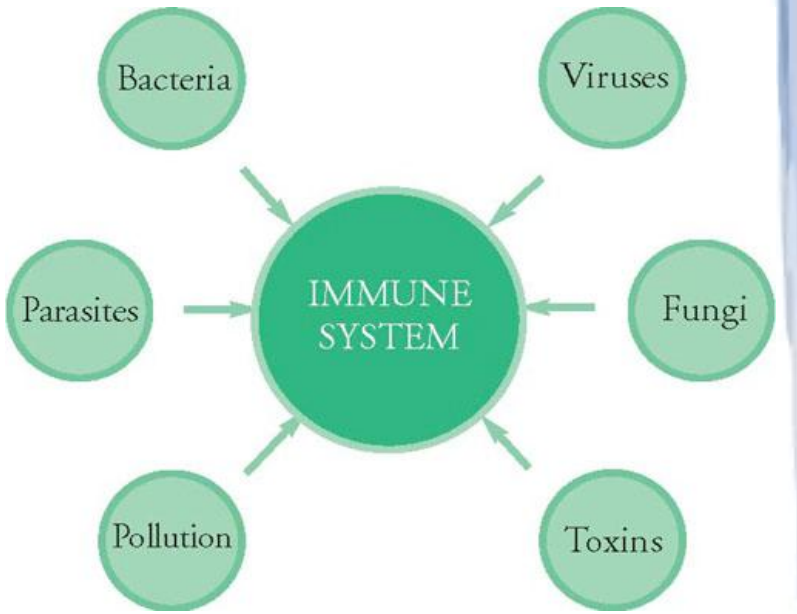
- Very similar to our legal system
- Store a large selection of cases
- Lookup engine
- Find case like problem at hand
- Example
  - The last time the car would not go, it had gas, and the battery was good, the problem was a plugged fuel filter.

Using Case Based Reasoning

- Must have library of cases
- Inference engine is hard to create, looking for similarities between problem and database of cases
- Cannot solve anything that was not in the original database
- Where is the intelligence?

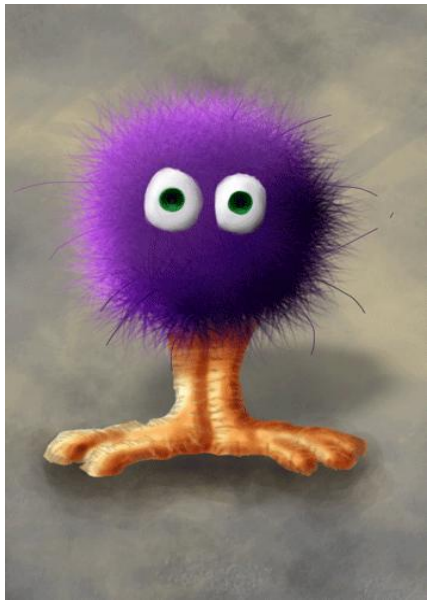
Synthetic Immune Systems

- Mimics human immune system
- Good for network security
- Detects intrusions
- Somewhat a reverse cluster analysis
- Detects if not in acceptable cluster
- Uses statistics, clustering, pattern recognition, etc
- Where is the intelligence?



Fuzzy Logic

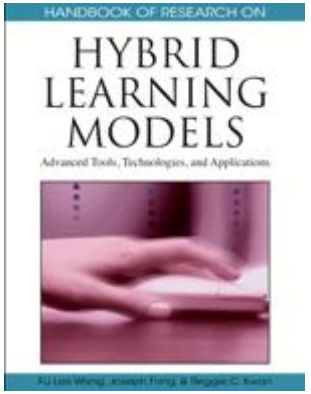
- Type I
  - is also known as PNL or Precisiated Natural Language
  - Lotfi Zadeh to communicate lingual variables – 1968
  - To convert lingual variables into computer inputs
  - High, medium, low, too much, etc
- Type II
  - Lotfi Zadeh & Jerry Mendel – 1975
  - Handles systems with uncertainty



Generative Adversarial Network

- Unsupervised learning
- Two learning systems
  - Opponents
  - Trying to create ‘fake’ to fool the other
  - Contest with each other
  - Both keep getting smarter
  - Creates high quality simulated object/image
- Checkpoint guard example
- Counterfeiter example
- Introduced NIPS 2014 by Goodfellow et al

Hybrids



- Combine several systems
  - GA and ANN
  - ANN with fuzzy, GA, & database
  - Many possibilities
- Uses more methods than just one type
- Can seed system with expert knowledge and then update with data
- Sometimes hard to get all parts to work together
- Ensemble learning classifier
- Harder to validate model
- Might require bigger training sample

Conclusion

- No cognitive process
- The intelligence is that of the programmer
- Biological intelligence is very complex
  - System of systems
  - We do not understand it all
  - Electrical AND chemical
- AI systems today attempt to emulate one facet of biological intelligence
- These AI systems can be very helpful

References

- The IEEE www.ieee.org
- www.machine-cognition.com (my website)
- American Association for Artificial Intelligence www.aaai.org
- IEEE Intelligent Systems Journal
- *Artificial Intelligence: A Modern Approach*, Russell & Norvig
- *Human Problem Solving*, Newell & Simon
- IEEE Computational Intelligence Society: www.ieee-cis.org

Future

- The Japanese have several personal servant robots. But are they intelligent?
- Augmented Cognition (AugCog)
  - Using AI methods to condense and combine in order to assist decision maker, not overwhelm (highly adaptive)
- My look into the future is that real intelligence is more elusive and complex than artificial intelligence
  - This has driven me to study biologically inspired computing
  - The EU has a new ruling requiring the ability to show why

Contact Information

Dr Brad Morantz
bradscientist@machine-cognition.com
http://www.machine-cognition.com
480-336-3400

Questions?