ENGINEERS AUSTRALIA

CPD Online

Machine Learning in (I)IoT

Dr John Ypsilantis MIEaust CPEng

Principal, Heuristics Australia Pty Ltd

Host – Michael Crapis

9 June 2020

ENGINEERS AUSTRALIA

Applied IoT
Community of Practice

ENGINEERS
AUSTRALIA

Engineering is a mixture of Science, Art and Black Magic

(H. Yee, c. 1984)

ENGINEERS
AUSTRALIA

The focus of this presentation is a presentation of basic concepts in AI/ML, and brief descriptions of semi-statistical and symbolic Machine Learning techniques which may be applied in (I)IoT contexts.

Mathematics are deliberately excluded (as far as possible)!

For (I)IoT practitioners, the intent is to present some ML options that may be useful in your day-to-day work but that you may not otherwise have considered

ENGINEERS
AUSTRALIA

Overview of this presentation

- What is an Artificially Intelligent (*AI*) System?
- What is Machine Learning (*ML*)?
- Classes of ML
- Some ML Algorithms
- ML in the context of (I)IoT
- Questions

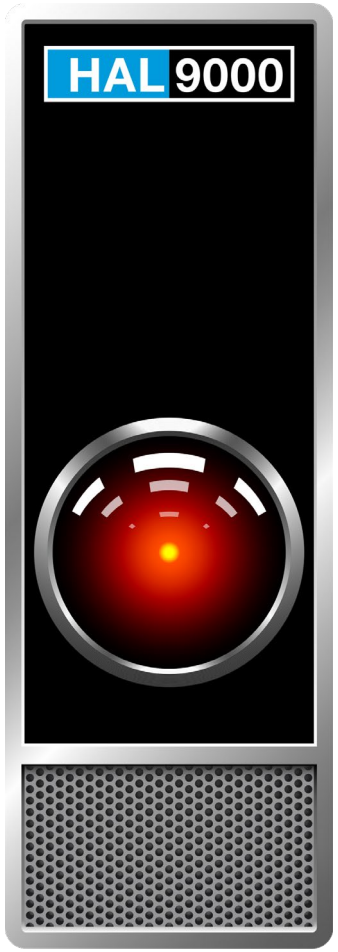
ENGINEERS
AUSTRALIA

Definitions for Artificial Intelligence and Machine Learning

ENGINEERS
AUSTRALIA

What is an Artificially Intelligent System?

(This is what typically comes to mind…)



By Grafiker61 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=46424786

ENGINEERS
AUSTRALIA

Definition

An *Artificially Intelligent System* is a software and/or hardware system which *approximates closure* of ("solves") an *N-P Hard* problem using *Heuristic Search Techniques*

Notes:

- *N-P Hard = intractable = cannot be solved definitively via an algorithm*
- *Heuristic Search Techniques = search methods and algorithms which are known to work well in most cases, but not necessarily for any formally provable reason*
- *"Solves" = creates at least reasonably good solutions most of the time*

ENGINEERS
AUSTRALIA

Machine Learning itself is a class of Artificially Intelligent System

Definition

A *Machine Learning System* is a software and/or hardware system which automatically produces descriptions of underlying regularities for a given process or set of examples using *Heuristic Search Techniques.*

ENGINEERS
AUSTRALIA

What is a Heuristic?

<u>Definition</u>

A *heuristic* is a *rule of thumb*, usually (but not always) without any scientific, mathematical or engineering theory to describe *why* it works

We all develop and eventually rely upon heuristics as we gain experience with a particular task or endeavor.

Heuristics are characteristic of *expert behaviour* – actions that are automatic and second-nature

ENGINEERS
AUSTRALIA

Some Background Concepts

ENGINEERS
AUSTRALIA

Characteristics of *intelligence*

- Learning
- Analogical Extrapolation
- (Heuristic) Search – “gut feel”
- Pattern Recognition
- Adaptation, Improvement
- Optimisation, Efficiency

→ *Intelligence*

(this is not an exhaustive list)

ENGINEERS
AUSTRALIA

Synonyms

AI/ML ⟷ *optimisation* ⟷ *search* ⟷ *pattern matching*

These relationships are ubiquitous within AI/ML, and can be considered fundamental

e.g., Backprop ANNs learn an *optimal* vector of synaptic weights to classify a set of *patterns*, by applying a heuristic *search* (gradient descent)

e.g., ID3 determines *optimal* decision points for analog variables in a vector, using a heuristic *search* based on the principle of entropy – (2nd law of thermodynamics or information theory)

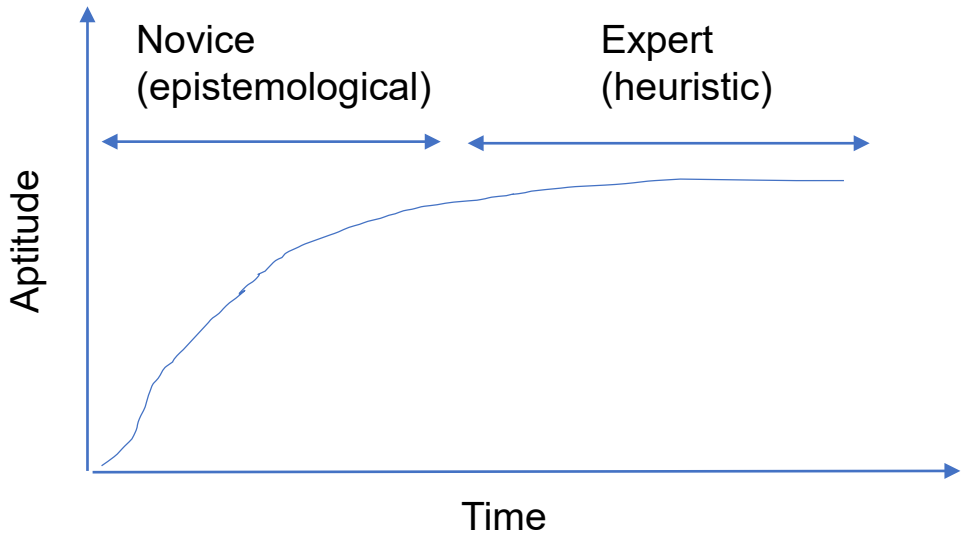
e.g., Genetic algorithms *search* for *optimal* solution vectors (*patterns*) by applying 3 types of heuristic *search* (replication, crossover and mutation)

ENGINEERS
AUSTRALIA

Martin's Law (W. H. Martin)

You can't learn anything unless you almost know it already

Learning Curve



ENGINEERS
AUSTRALIA

Some Examples of Heuristic Search

ENGINEERS
AUSTRALIA

What is Heuristic Search?

<u>Definition</u>

Heuristic Search is a class of search techniques that make use of *heuristics* rather than conventional computational algorithms

Heuristic search very often finds novel, “better” or globally optimal solutions, as there is little or no reliance on assumptions such as linearisation or other simplification of the problem space (which in turn often distort the problem that is being solved in the first place)

The key disadvantage is that heuristic search often requires significant time (large number of iterations) to locate a “good” solution

ENGINEERS
AUSTRALIA

Some Examples of Heuristic Search in Machine Learning

- *Gradient Ascent/Descent*
- *Swarm Algorithms*
- *Genetic Algorithms*
- *Simulated Annealing*

ENGINEERS
AUSTRALIA

Gradient Descent/Ascent

A minimum/maximum is likely to be in the direction of maximal gradient

ENGINEERS
AUSTRALIA

Swarms

The best solution is likely located at a point toward which the majority of independent agents (in the natural world, bees, birds, fish etc) are moving

ENGINEERS
AUSTRALIA

Genetic Algorithms

The best solution is likely found by allowing a population to be dominated by relatively good (strong) solutions (*reproduction*). This is augmented by synthesizing candidate solutions by cut-and-paste (*crossover*) and random modifications of individual elements (*mutation*)

ENGINEERS
AUSTRALIA

Simulated Annealing

The best solution is likely to be found by randomly searching and creating a finite number of clustered ”good” solutions, as “temperature” is gradually reduced –
analogous to recrystallisation by annealing (movement to a lower energy state)

ENGINEERS
AUSTRALIA

Why do heuristics work?

We don't really know.

However, as they regularly occur in nature
they are likely to be very good, possibly even optimal.

Biological paradigms in particular are the result of
over 4 billion years of continual evolutionary improvement
and optimisation, so it would be expected that
they would work well!

ENGINEERS
AUSTRALIA

Classification of Machine Learning Algorithms

ENGINEERS
AUSTRALIA

Classification of Machine Learning Algorithms

Autonomy

Supervised

Requires *tagged* training vectors

e.g., *X* is characterized by (*y,z,w,t,…)*
or
(y,z,w,t,…) is an example *of X*

Unsupervised

Can work with *untagged* training vectors

e.g., (*y,z,w,t,…),*
(a,e,q,f,…),
etc…

ENGINEERS
AUSTRALIA

Classification of Machine Learning Algorithms

Learning Mechanism

Statistical

Rely purely on statistical processes

Examples:

- *Bayesian Inference*
- *Hidden Markov models*

Semi Statistical

Employ stochastic heuristics

Examples:

- *ID3, C4.5, C5.0*
- *Artificial Neural Nets (backprop)*

Symbolic

Employ relational properties and heuristics

Examples:

- *Version Spaces*
- *Semantic Nets*

ENGINEERS
AUSTRALIA

Characteristics of Statistical “Learning”

Statistical methods require **large** datasets for efficacy

In contrast, Semi-Statistical and Symbolic methods can often learn effectively from small datasets, sometimes as few as 2 or 3 examples.

Many applications are well suited to statistical learning and inference, particularly in the big data/IoT space

<u>However, consider this</u>: how many instances of experiencing a life threatening situation will result in effective learning by a human? One or two, or the order of tens or hundreds?

Semi-Statistical methods in particular apply well in many (I)IoT contexts, and better leverage the simulation of true learning and intelligence

ENGINEERS
AUSTRALIA

Some Examples of Machine Learning Algorithms

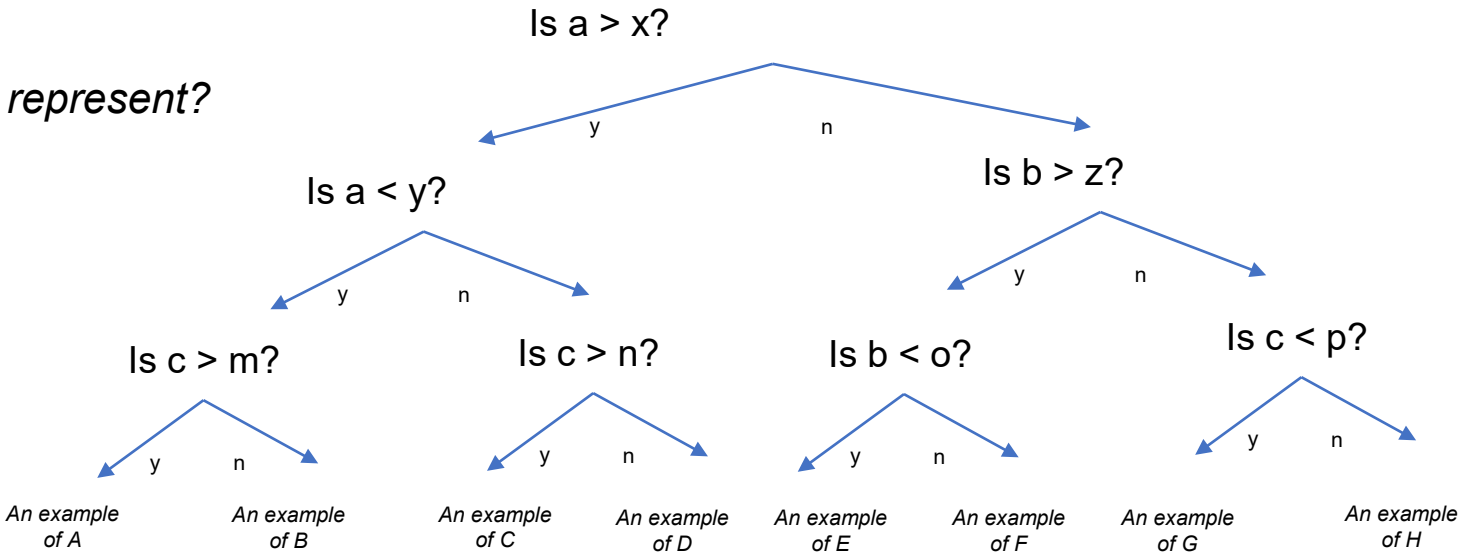
ENGINEERS
AUSTRALIA

Example – ID3

- ID3 is a Supervised, Semi-Statistical Learning Algorithm
- It accepts tagged examples, comprising multiple classes, e.g., vectors of blood test results, tagged with manually diagnosed diseases.
- It produces an efficient (binary) *decision tree*, a binary tree where each node comprises a test over a vector that is to be diagnosed
- ID3 selects tests based on maximal reduction in total entropy of the two resulting subsets (*partitions*) of vectors.
- It proceeds recursively until a termination condition is reached
- The leaves of the tree each correspond to a classification or diagnosis

Example – ID3

What does (a,b,c) represent?



ENGINEERS
AUSTRALIA

Example – ID3

Potential applications to IIoT:

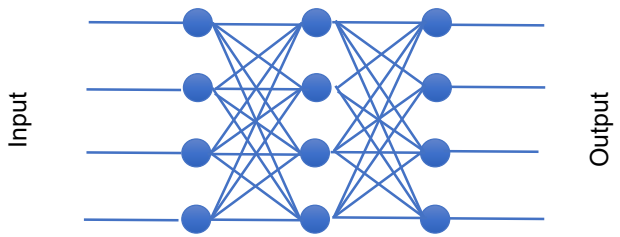
- Elicitation of diagnostic information from representative sets of sensor data vectors
- Subsequent driving of responses to changes in sensed state: *analysis-response cycles*

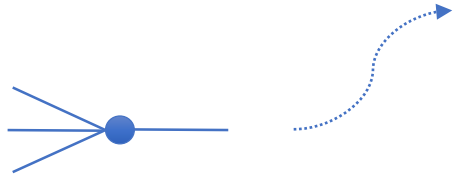
ENGINEERS
AUSTRALIA

Example – Artificial Neural Networks (back propagation)

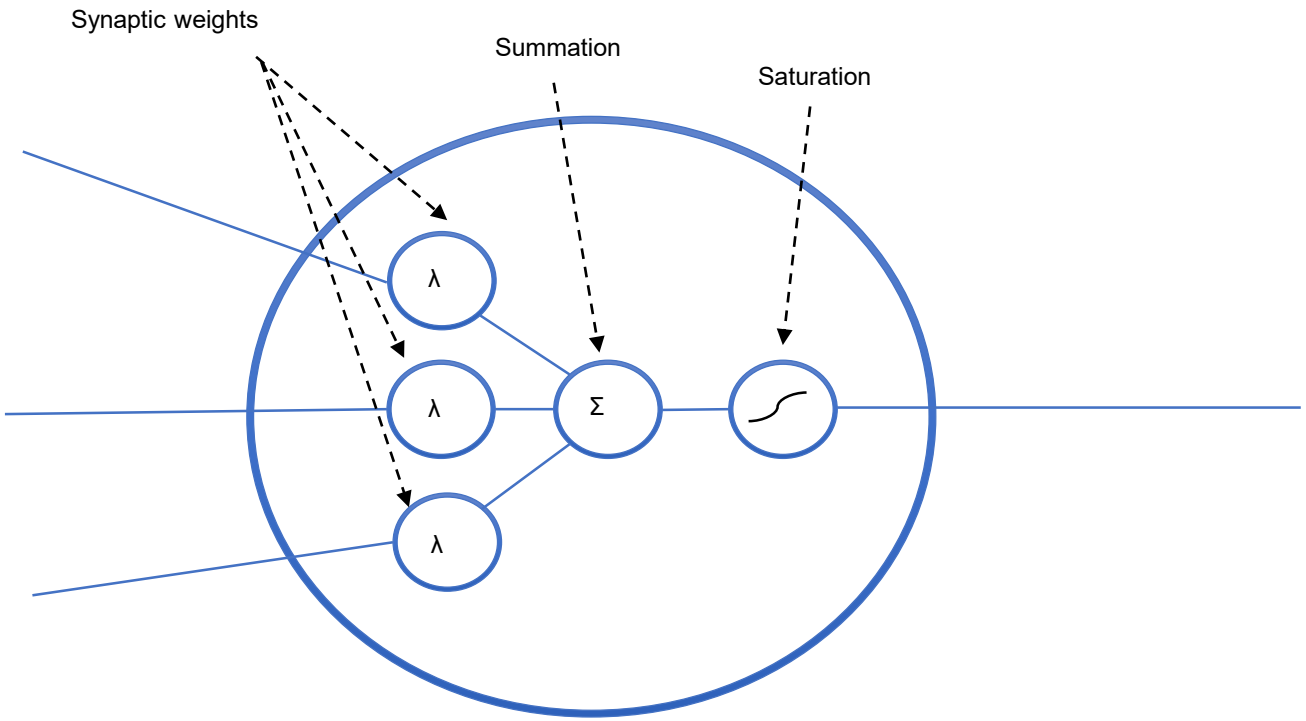
- Backpropagation ANNs employ a Supervised, Semi-Statistical Learning Algorithm
- The ANN accepts tagged examples of a single class, e.g., (noisy) vectors corresponding to a specific pattern, such as an encoded alphabetic character, ECG etc.
- The encoded tag is introduced to the output of the ANN, the vector to the input, and *synaptic weights* are adjusted to match input to output via a *gradient descent* heuristic
- ANNs can efficiently recognize patterns for which they have been trained, even when the input is noisy or somewhat different to any of the training examples

Example – Artificial Neural Networks (back propagation)





A typical *neuron* (basic computational element)



ENGINEERS
AUSTRALIA

Example – Artificial Neural Networks (back propagation)

Potential applications to IIoT:

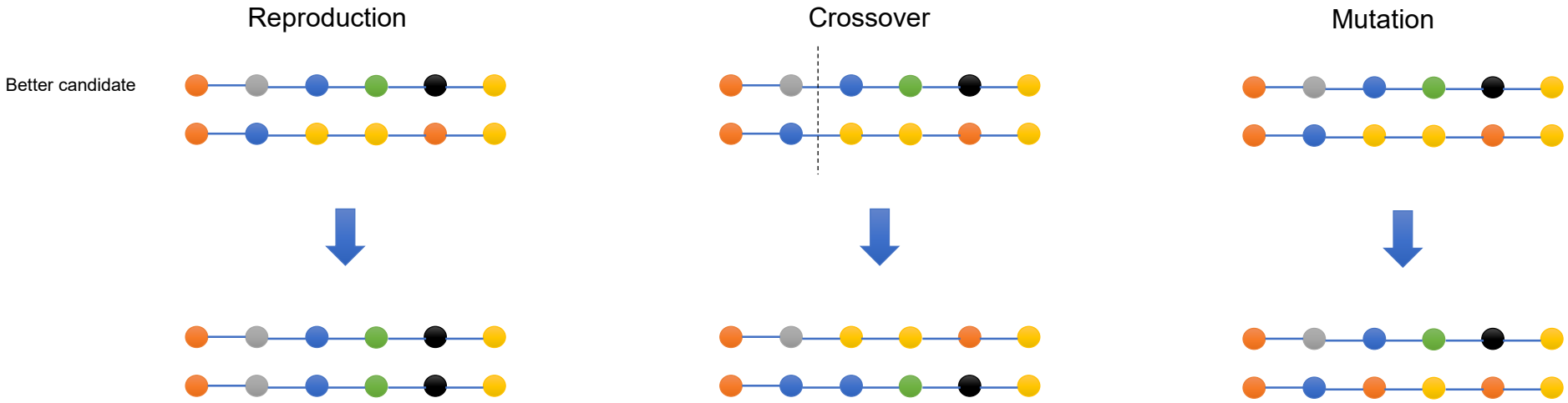
- Elicitation of diagnostic information from large sets of sensor data, particularly for identification of patterns in such data
- Subsequent driving of responses to changes in sensed state: *analysis-response cycles*

ENGINEERS
AUSTRALIA

Example – Genetic Algorithms

- Genetic Algorithms employ Semi Statistical Heuristic Search to find an optimal solution vector, i.e., a vector that maximises a predetermined *objective function*, by generating populations of candidate solutions
- These algorithms perform three basic operations:
    1. *Reproduction:* replication of good solutions at the expense of “less good” solutions
    2. *Crossover*: random cut-and-splice between pairs of solutions, to insert novel candidates, and
    3. *Mutation*: Occasional, minor perturbation of individual elements of vectors, to insert novel candidate solutions (that may potentially “break out” of local minima/maxima)

Example – Genetic Algorithms



ENGINEERS
AUSTRALIA

Example – Genetic Algorithms

Potential applications to IIoT:

- Optimisation over large datasets with nonlinear objective functions, e.g. determining the optimal delivery sequence for a group of couriers, or optimal dispatch of work crews for a utility

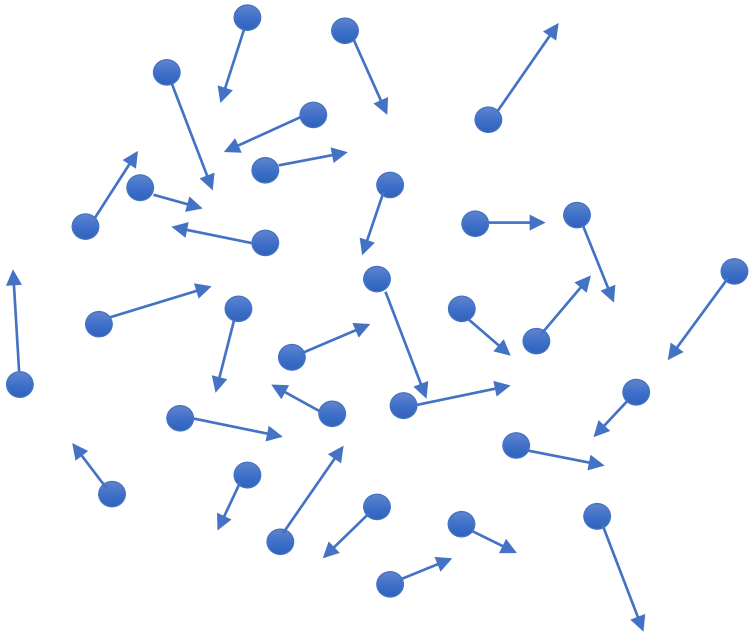
ENGINEERS
AUSTRALIA

Example – Simulated Annealing

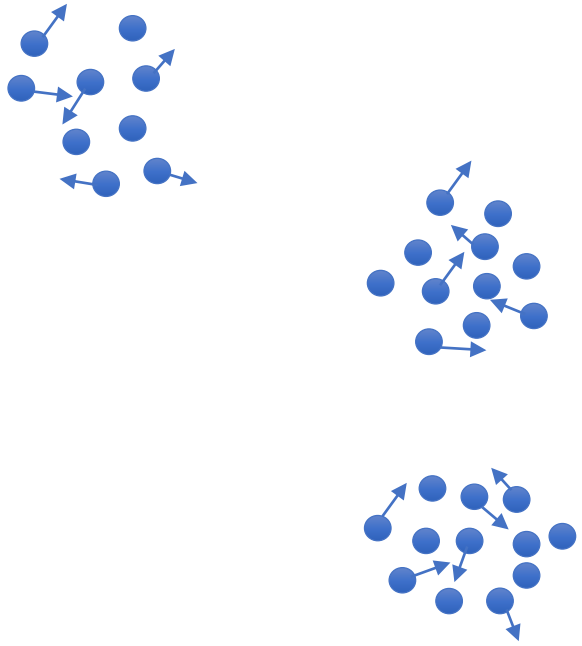
- Simulated Annealing employs a Semi Statistical Heuristic Search to find an optimal solution vector by generating clusters of “good” candidate solutions (*particles*) using a thermodynamic paradigm (*entropy*)
- Candidate solutions are perturbed in random directions, with velocities proportional to a *temperature* parameter
- The temperature parameter is progressively lowered over time, with resultant reduction of the velocity and distance travelled by the particles
- Particles which evaluate to higher scores (via the function that is being optimized) move with inversely proportional speed at a given temperature
- Eventually, clusters of particles will form. Each comprises a relatively “good” solution to the optimization. The optimal value is the centroid with the highest score.

Example – Simulated Annealing

High temperature



Low temperature



ENGINEERS
AUSTRALIA

Example – Simulated Annealing

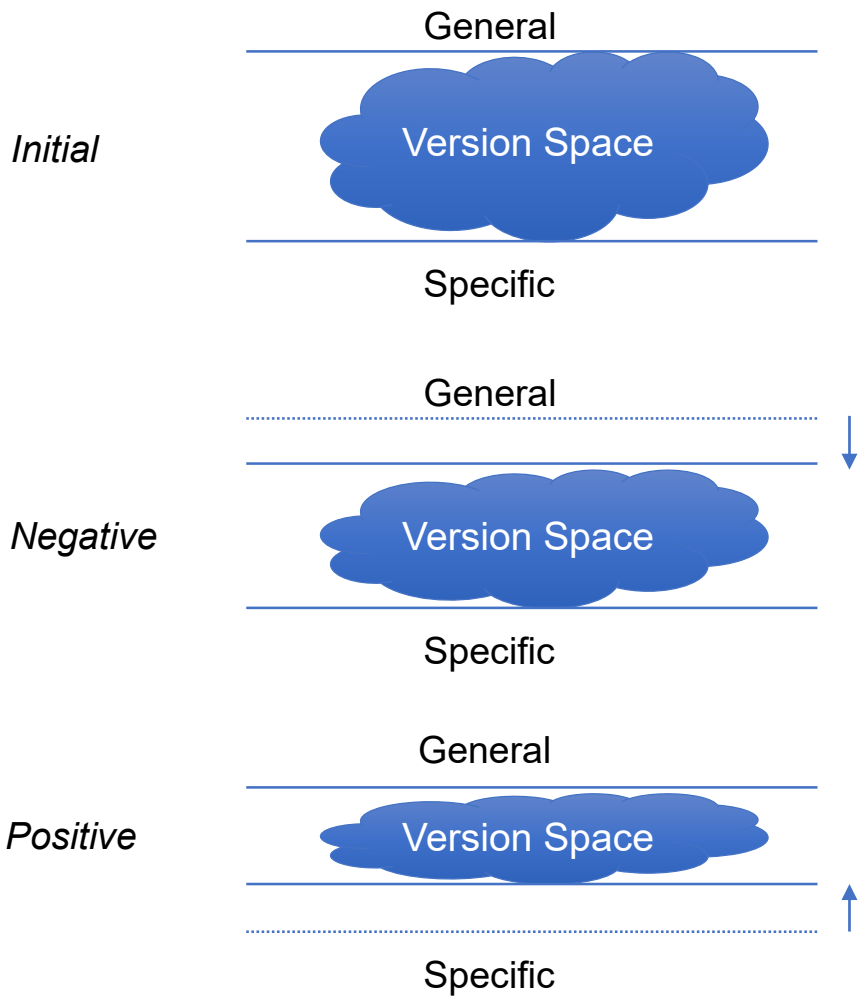
Potential applications to IIoT:

- Optimisation using large datasets and/or nonlinear objective functions

ENGINEERS AUSTRALIA

Other machine learning algorithms – Version Spaces

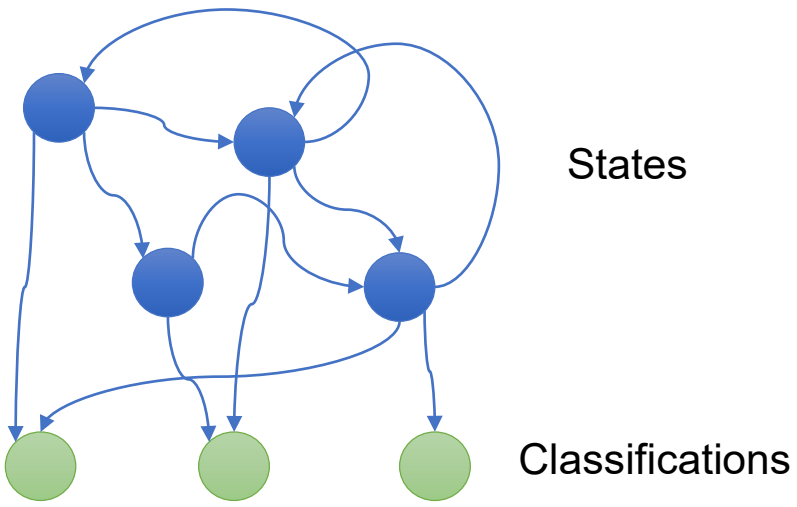
- Supervised Symbolic learning algorithm
- Defines a (large) space of candidate solutions with a *General* and a *Specific Boundary*, to reduce computational complexity
- Positive examples *prune* the space by moving the Specific boundary (*generalization heuristic*)
- Negative (Counter) examples prune by moving the General boundary (*specialization heuristic*)
- Accuracy is progressively improved with new examples
- Learning is complete when the boundaries *converge*
- IIoT applications include elicitation of diagnostic information from tagged examples



ENGINEERS
AUSTRALIA

Other Machine Learning Algorithms – Hidden Markov Models

- Supervised, Statistical learning algorithm
- Generates a stochastic *state machine*, with probabilities for each state transition, and probabilities for each outcome
- HMMs can be applied to (I)IoT scenarios where it is necessary to determine and identify *sequence* as well as pattern
- IIoT applications include elicitation of diagnostic information which relies on sequence, e.g., signature analysis

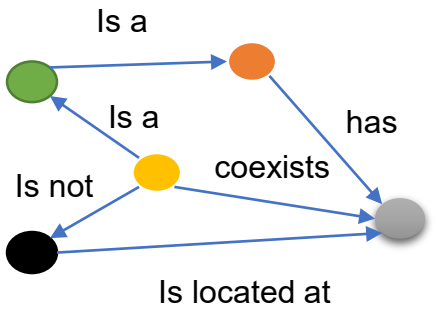


Each arc has a probability of transition

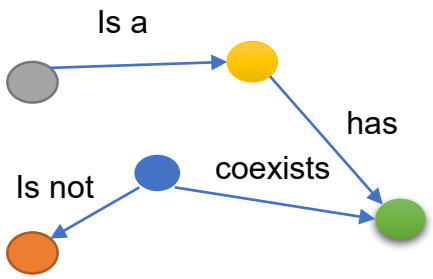
ENGINEERS AUSTRALIA

Other Machine Learning Algorithms – Semantic Networks

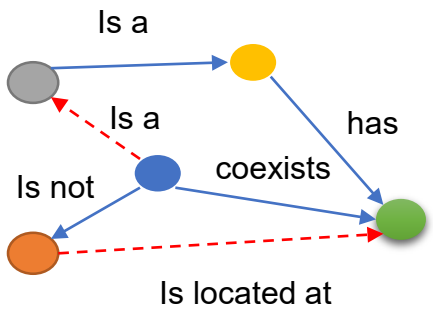
- Supervised, Symbolic learning algorithm
- Generates a graph of *relations* over a given domain
- Can be used to postulate relations over a different domain, by analogy
- IIoT applications include discovery of new relationships on the basis of previously seen examples, over multiple domains (analysis of data)



Context A



Context B



Analogical Inference From Context A

ENGINEERS
AUSTRALIA

Hypothetical Examples of Machine Learning Applications in IIoT contexts

ENGINEERS
AUSTRALIA

Location of damaged utility infrastructure

- For example, the location of damaged points in a gas or water network using ubiquitous IIoT telemetry and sensing
  - A pattern classifier such as an ANN can be trained on fault data generated via a simulator
  - The ANN can subsequently analyse real-time data to predict the presence and location of a “hit”

ENGINEERS
AUSTRALIA

Dispatch of a fleet of service vehicles

- For example, dispatch of utility work crews in an environment where calls and issues arise in real time during the course of the day, requiring continual rescheduling and reprioritization of work
    - An optimiser based on a genetic algorithm or simulated annealing algorithm, can continually optimize a constrained objective function to minimize time, cost and response time for a fleet of service vehicles, taking their current location into account

ENGINEERS
AUSTRALIA

Condition-Based Asset Management

- For example, monitoring and subsequent servicing of machinery on the basis of condition rather than merely operating time
  - Use an algorithm such as a Sequence Graph or Hidden Markov Model to identify and classify signatures from IIoT sensing
  - Subsequently, use an optimiser based on a genetic algorithm or simulated annealing algorithm to determine an optimal scheduling of maintenance, taking into account the minimization of cost, minimization of outage time and maximization of equipment life

ENGINEERS
AUSTRALIA

Fall Detection in Aged Care Facilities

- Need for appropriate monitoring to ensure resident safety, but with regard to privacy and dignity
- Sensing must therefore be relatively simple
    - Use an algorithm such as a Sequence Graph or Hidden Markov Model, in conjunction with a pattern matching front end such as an ANN, to identify and classify signatures from IIoT sensing corresponding to a fall. An automatic alarm may then be raised to staff

ENGINEERS
AUSTRALIA

Thank you!

For more information:
Email: john.ypsilantis@heuristics.com.au

http://www.heuristics.com.au

ENGINEERS AUSTRALIA

CPD Online

Thank you

Today's presenter:

Dr John Ypsilantis –
john.ypsilantis@heuristics.com.au

Join the community

https://iot.engineersaustralia.org.au

Applied IoT
Community of Practice