How effective are machine learning algorithms compared with traditional analytical techniques, with respect to playing abstract games?

Edmund Goodman – L6 REL
Rouse Research Essay Award
May 2019

Figure 1: A piece of abstract art, generated from an input image of a chess set [1], using a convolutional neural network called DeepDream [2] to "find and enhance patterns in images via algorithmic pareidolia"

Word count: 4030

Le défaut unique de tous les ouvrages
c'est d'être trop longs.
— VAUVENARGUES, *Réflexions*, 628 (1746)

Figure 2: Donald Knuth: The Art of Programming Volume 1: Fundamental Algorithms, page xiv [33]
Translation: "The only defect of all works is to be too long"

My essay is in the subject area of computer science, and my project supervisor was Mr Gwilt, who I would like to thank for many interesting conversations and comments throughout the whole project.

0 Abstract

Machine learning is an increasingly commonly used technique for solving many problems, and in some cases replacing other analytical techniques which were the previous status quo.

In this essay, we investigate the effectiveness of both machine learning algorithms and analytical techniques, with respect to playing abstract games. We first describe neural networks from first principles as a tool for problem solving, considering their biological origins, structure, and processes for training them. We achieve this through a combination of research, mathematics and programmatic representations. Next, we consider analytical techniques, and provide the examples of mathematical analysis, genetic algorithms and tree search approaches, again researching techniques, then implementing the ourselves. During this phase, we found that we yielded different results to those reported by popular mathematician Martin Gardner in his column in the Scientific American, which highlighted a subtle difference in training routines. Following this, we consider the cases where these techniques have been applied to games, taking the examples of AlphaGo, Stockfish and Chinook, and seeing the development of these algorithms from early predictions to the cutting edge.

We found that machine learning techniques can be very effective when applied to playing abstract games, and have been shown to be able to beat traditional analytical techniques in some cases. However, they are not a silver bullet, and should not be applied to many problems. For example, very simple games, where analytical techniques require less processing time, and yield better results. This result has far-reaching consequences in almost every field of research, abstract games are effective analogues for many problems, and a new, very effective, technique to solve them could help many researchers.

1 Table of Contents

2 Introduction

The Merriam Webster dictionary defines games as “a physical or mental competition conducted according to rules with the participants in direct opposition to each other” [3]. They have been a staple of human society since the earliest civilisations, with the earliest complete board game “The Royal Game of Ur”, dating back to 2500 BC [4], and the net worth of the gaming industry in 2018 was $43.8 billion [5]. Abstract games are a specific type of game, in which the theme is unimportant to the experience of gameplay, and hence can be usefully analysed by computer algorithms. They tend to be combinatorial, and played by two parties alternating a finite number of terms.

Abstract games are a useful place to develop techniques for solving problems in computer science, as these techniques can often be transferred to “real life” [6], for example, game theory is very prevalent in economics, and neural netwo

rks are used to perform microtransactions generating huge amounts of revenue [7].

As a result of this considering algorithms to play abstract games is immensely useful, since many techniques developed can be applied to real world problems. Furthermore, they are more accessible to people without a grounding in STEM, as it is easier to consider a game of tic-tac-toe than classifying a dataset of 60,000 handwritten digits [8], despite the fact both problems might have similar solution vectors.

Machine learning is a very popular issue [9], and it is being used in many areas of research [10] [11] [12]. However, it is important to contrast the current surge in development of machine learning to other already developed techniques that might be more suited to a specific problem.

In summary, abstract games are a great place to discover and develop algorithms which are useful in the real world, and both machine learning and traditional analytical techniques can be used to solve them.

3 Neural networks from first principles

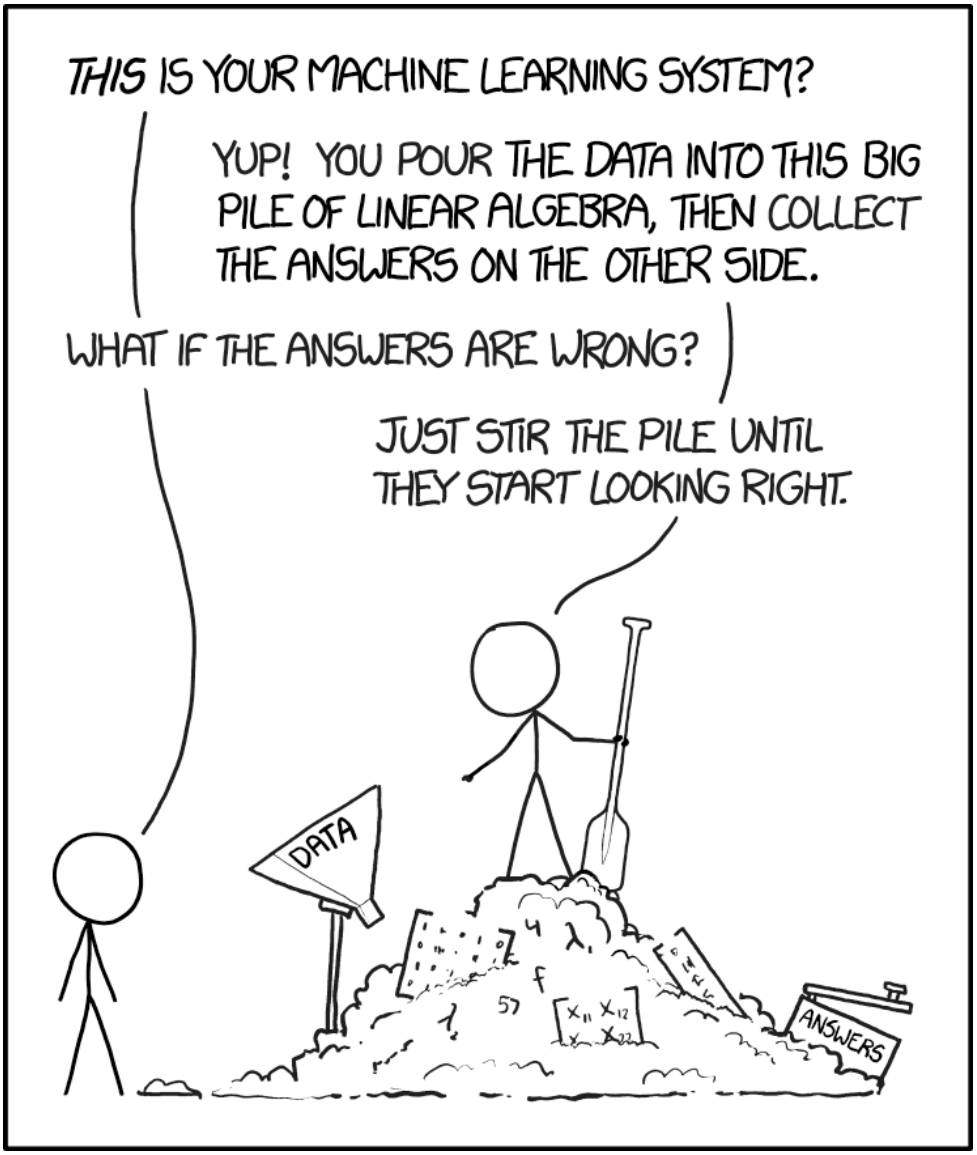


Figure 3: An XKCD comic by Randall Munroe addressing modern machine learning [13]

3.1 Introduction

Neural networks allow machines to learn from data without human intelligence, and in recent years have gained prominence as a means of solving many problems where other techniques have failed, or can only be applied to specific subsets of the problem space. They are normally used to cluster or classify large sets of data, and can often identify features other techniques and humans wouldn't find, for reasons including: a very large or multidimensional search space; or the features being very difficult to perceive manually. However, they can also be applied to many other tasks, albeit with varying effectiveness [14].

Neural networks can be loosely considered as a mathematical model for processes that occur in biological brains, and their underlying structure was based off human neural structures.

3.2 From biological to mathematical systems:

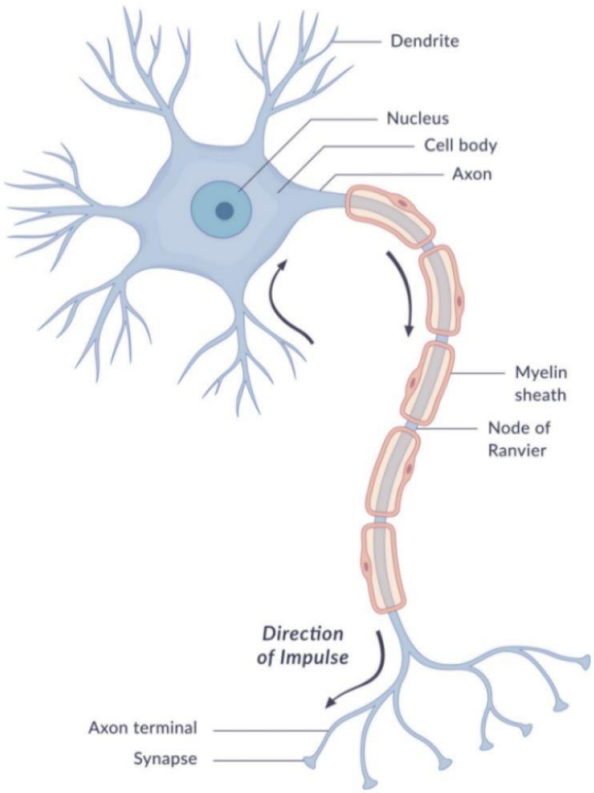


Figure 4: The structure of a human neuron [15]

As shown in Figure 4 (above), dendrites collect input signals from previous neurons' synapses, then the combined effect of all these signals passes into the axon. This signal then passes out into multiple axon terminals, and into then into synapses, which passes signals into other dendrites, and the process repeats itself [16]. This system appears simple, but the cumulative effect of billions of interconnected neurons is the definition of human intelligence.

We can produce a mathematical model for this system. We call the neuron a node, which has inputs, like the dendrite, an activation function based on the cumulative effect of all the inputs, like the synapse, and outputs, like the axon terminal.

This mathematical model for the biological system was first proposed by Frank Rosenblatt as "Perceptrons" [17] [18], but has been developed over time.

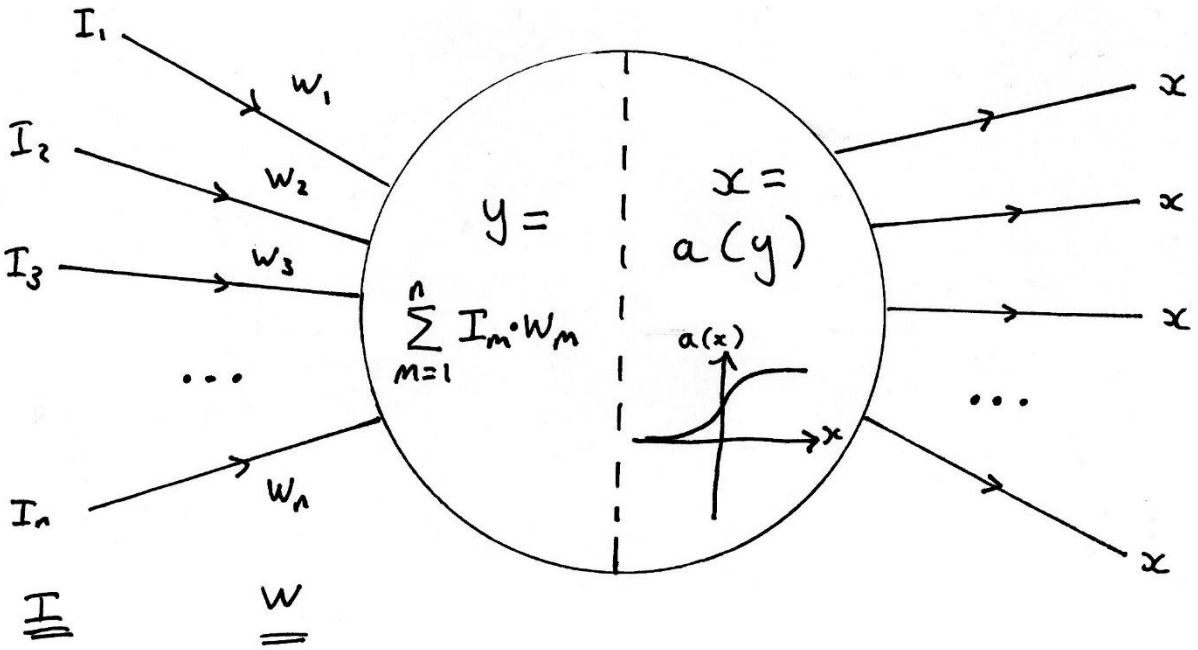


Figure 5: A diagram of a single node in a network

The output value of each node is evaluated as the sum of the products of its input values, I, and their weights passed, W, through an activation function, $a(x)$:

$$x = a\left(\sum_{m=1}^{n} I_m W_m\right)$$

A ubiquitous activation function is the logistic sigmoid activation function [19], which was one of the first to be used, as early computer scientists thought that its mix of a central pseudolinear section and curved edges would yield good results.

$$a(x) = \frac{1}{1 + e^{-x}}$$

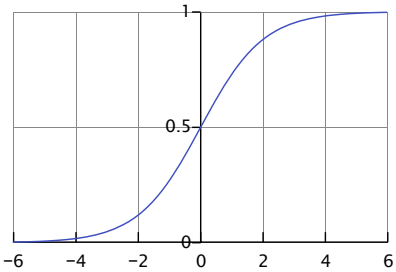


Figure 6: A sigmoid curve [62]

It has stood the test of time, but other functions are also used, as detailed in appendix 7.6 (page 39)

3.3 Building a network

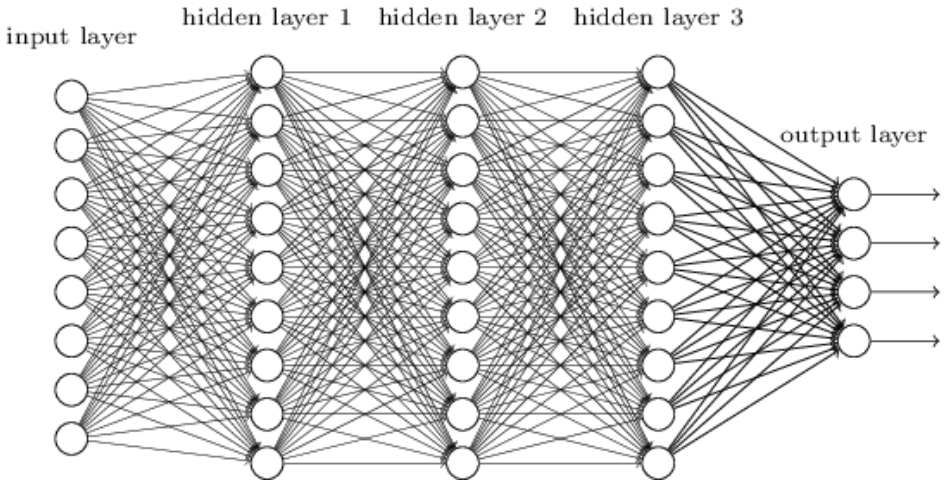


Figure 7: A deep neural network [20]

It is clear to see that individual nodes are not especially useful owing to their simplicity. However, we can link them together to form networks, of increasing complexity, which hence allows more complex data to be modelled.

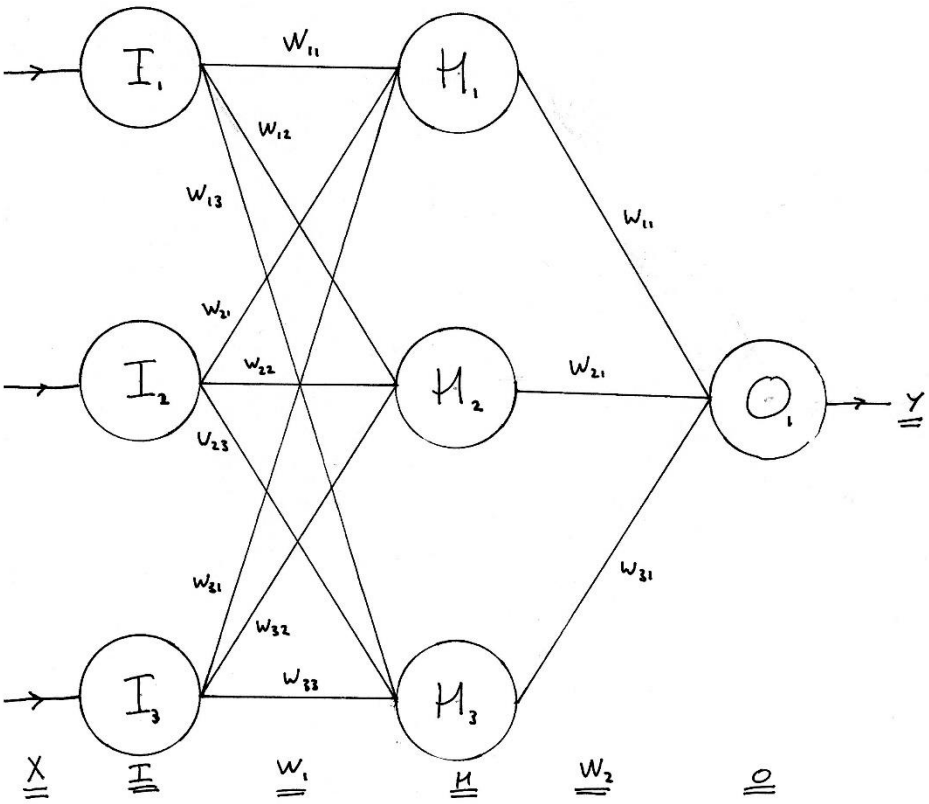


Figure 8: A small neural network, labelled to denote nodes and weights

Traditionally, a network is modelled as having input nodes, where each input data point maps directly to one node. Next, there are hidden layers, which for simplicity we will consider as shallow and fully connected, as shown in the Figure 8 (above). Each node in the hidden layer is connected to every node in the previous layer, and each of those

connections have a weight. Finally, we have output nodes, which have many inputs, but only produce on output. This model means we can put data into the network via its input nodes, and collect its output from the output nodes. The process of putting data though the network is known as forward propagation, and is performed by iterating over each layer, and for each node in that layer, evaluating it as previously described, and yielding its outputs. This process is known as forward propagation.

$$H_n = a\left(\sum_{m=1}^{3} I_m W_{nm}\right)$$

$$O = a\left(\sum_{m=1}^{3} H_m W_{nm}\right)$$

This can also be expressed as a matrix multiplication, which is preferable during implementation, as there exist heavily optimised libraries to perform matrix operations, e.g. NumPy for python

$$H = a(I \cdot W)$$

$$O = a(H \cdot W)$$

I implemented forward propagation in python, to ensure I fully understood it. A screenshot of the basic python code is shown in Figure 9 (below):

```
def forwardpropagate(self, X):
    """Forward propagate an input vector X through the network composed of
    the weights self.W1 and self.W2, storing intermediary variables self.z
    and self.a to allow back propagation, and returning the output vector
    resulting from the process
    """

    #Forward propagate data though the network
    self.z, self.a = [], []
    #From input to hidden layer
    self.z.append( np.dot(X, self.W1) )
    self.a.append( self.activation(self.z[-1]) )
    #Through the hidden layers
    for i in range(len(self.W2) - 1):
        self.z.append( np.dot(self.a[-1], self.W2[i]) )
        self.a.append( self.activation(self.z[-1]) )
    #From hidden to output layer
    self.z.append( np.dot(self.a[-1], self.W2[-1]) )
    self.a.append( self.activation(self.z[-1]) )
    #Return the final value of the output layer
    return self.a[-1]
```

Figure 9: A python implementation of forward propagation

In the very early days of development, networks were massive circuits, with each weight being a potentiometer that was manually adjusted [21] [22]. However later they became entirely computationally modelled, which allowed for more effective training algorithms.

Now, we have described a network we need to be able to teach it to produce the output data we want from the input data we want. In order to train a network, we take our input data, and forward propagate it through our network, and then compare it to our expected output data. This comparison of the expected against the calculated yields our networks error, which we want to minimise to train our network. A naïve way to approach this would be to randomly try different weights, until we find a set that is good enough for our use case, however, this is very inefficient.

3.4 Training and the backpropagation algorithm

The backpropagation algorithm [23] [24] is a way to train networks, by considering how much each node contributed to the error in the output and adjusting the weights accordingly.

It transpires there are many online tutorials detailing single layer backpropagation, but multilayer backpropagation tends not to be covered, resulting in me working deriving much of the above mathematics myself.

The mathematical derivation of the proof is immensely complicated [25], and far out of scope, however, we can summarise the final equations:

First, we will define the variables we will use:

$$y = the\ expected\ output\ of\ the\ network$$
$$\hat{y} = the\ output\ of\ forward\ propagating\ the\ network$$
$$X = the\ input\ to\ the\ network$$
$$H_n = the\ value\ during\ forward\ propagtion\ in\ the\ n^{th} hidden\ layer$$

$$W_{\overrightarrow{H_n O}} = the\ weight\ matrix\ between\ the\ final\ hidden\ layer\ and\ the\ output\ layer$$
$$W_{\overrightarrow{H_n H_{n+1}}} = the\ weight\ matrix\ between\ the\ hidden\ layers$$
$$W_{\overrightarrow{I H_1}} = the\ weight\ matrix\ between\ the\ input\ layer\ and\ first\ hidden\ layer$$

$$E_n = the\ matrix\ containing\ the\ error\ in\ the\ n^{th}\ layer\ weights$$
$$\Delta_n = the\ matrix\ the\ change\ to\ be\ applied\ to\ the\ n^{th}\ layer\ weights$$

$$a'(x) = a(x)\big(1 - a(x)\big) = the\ derivative\ of\ the\ sigmoid\ activation\ function$$

$\eta = the\ learning\ rate\ of\ the\ network\ (a\ scale\ factor\ for\ the\ weight\ updates)$

And specifying ambiguous operands:
$X^T = the\ matrix\ transpose\ of\ X$
$A \circ B = the\ entrywise\ product\ of\ matrices\ A\ and\ B\ (hadamard\ product)$
$A \cdot B = the\ matrix\ product\ A\ \&\ B$

In essence, backpropagation is the process of tracing a route back through the network, considering how each weight contributed to the output error (the difference between the expected and calculated output values), then calculating a gradient to find which direction, and by how much each weight should be changed, and finally adjusting that based on the amount it contributed to the error.

First, we backpropagate from the output layer to the final hidden layer:

$$E_1 = y - \hat{y}$$
$$\Delta_1 = E_1 \circ a'(W_{\overrightarrow{H_2O}})$$

Next, we backpropagate within hidden layers (this step can be repeated for more hidden layers):

$$E_2 = \Delta_1 \cdot (W_{\overrightarrow{H_2O}})^T$$
$$\Delta_2 = E_2 \circ a'(W_{\overrightarrow{H_1H_2}})$$

Finally, we backpropagate from the first hidden layer to the input layer:

$$E_3 = \Delta_2 \cdot W_{\overrightarrow{H_1H_2}}^T$$
$$\Delta_2 = E_3 \circ a'(W_{\overrightarrow{IH_1}})$$

Then we can update all the weights, scaled by a learning rate η

$$W_{\overrightarrow{H_2O}} \mathrel{+}= \eta\left({H_2}^T \cdot \Delta_1\right)$$
$$W_{\overrightarrow{H_1H_2}} \mathrel{+}= \eta\left({H_1}^T \cdot \Delta_2\right)$$
$$W_{\overrightarrow{IH_1}} \mathrel{+}= \eta(X^T \cdot \Delta_3)$$

In order to train a network, we can repeat for cycle of forward propagation, back propagation and updating weights, known as epochs, many times, until the network models the input data set effectively.

As before, I implemented the processes of backpropagation: Figure 10, and updating weights: Figure 11, to ensure I fully understood them.

```python
def backwardpropagate(self, X, y, yHat):
    """Back propagate the error of the most recent forward propagation
    through the network, producing the arrays self.error and self.delta
    to allow weight updating
    """

    #Back propagate the error through the network
    self.error = []
    self.delta = []
    #Back propagate from output to the last hidden layer
    self.error.append( y - yHat )
    self.delta.append( self.error[0]*self.activation(yHat, True) )
    #Back propagate through the hidden layers
    for i in reversed(range(0, len(self.W2))):
        self.error.append( np.dot(self.delta[-1], self.W2[i].T) )
        self.delta.append( self.error[-1]*self.activation(self.a[i], True) )
```

Figure 10: A python implementation of the backpropagation of error through a network

```python
def updateWeights(self, X, learnRate):
    """Update the weights of the network based on the previous pass of
    backpropagation in order to improve the network model
    """

    #Update the hidden layer weights
    j = 0
    for i in reversed(range(len(self.W2))):
        self.W2[i] += np.dot(self.a[i].T, self.delta[j]) * learnRate
        j += 1
    #Update the input layer weights
    self.W1 += np.dot(X.T, self.delta[j]) * learnRate
```

Figure 11: A python implementation of the updating of weights following a pass of backpropagation

However, this training process has a few subtle modes of failure. Firstly, if the training set isn't large enough to sufficiently describe the data, or the network architecture is too simple, it doesn't produce a complete model of the data, which is called *underfitting*. Secondly, if the network is trained for too long on the data, it can produce a model that only fits the training data, not the general system, which is called *overfitting*. Finally, irrespective of good design and data, training networks requires time and processing power, and a major flaw of large networks modelling complicated systems is that they require many resources to train effectively.

There are various techniques which can be employed to address these issues, which we detail in appendix 7.6 (page 36), and an example training cycle on a simple data set in appendix 7.1 (page 24).

3.5 Summary

In order to fully understand how neural networks function, it is useful to implement them yourself, as the process of development and debugging ensures a full understanding and exposes subtle errors that might otherwise go unnoticed. There are also resources that provide a taste of what neural networks do, such as Tensorflow playground [26], shown in Figure 12 (below).

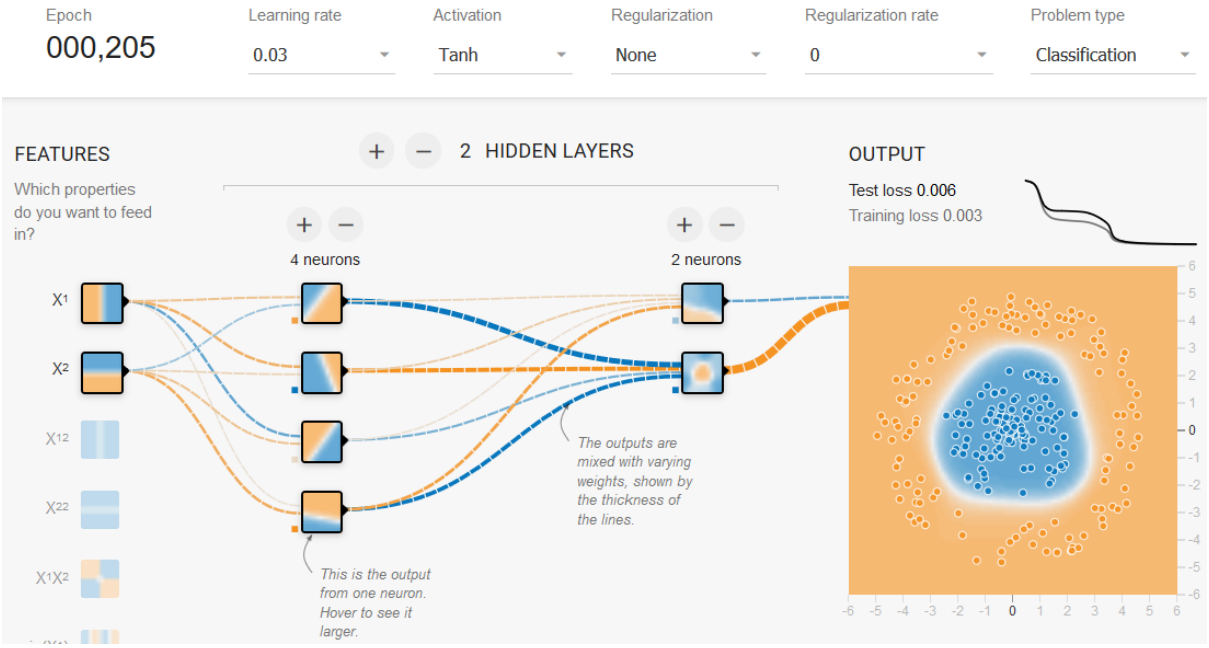
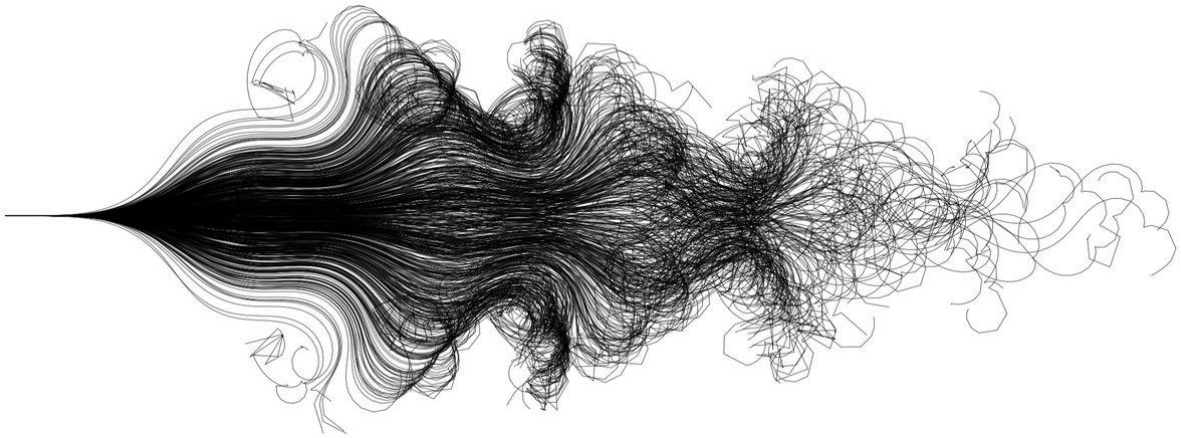


Figure 12: A screenshot of the Tensorflow playground

Neural networks are an increasingly ubiquitous technique in many fields of research, and many researchers are trying to, often naïvely, use them as a silver bullet for difficult problems. Generally, they only work on problems with a large, comprehensive training set, and are only useful if traditional analysis of that set is not viable. Nonetheless, they can be incredibly effective for some problem types, and will no doubt grow more important in the future.

4 Analytical techniques



Instability of an unsteered bicycle. This shows 800 runs of a bicycle being pushed to the right. For each run, the path of the front wheel on the ground is shown until the bicycle has fallen over. The unstable oscillatory nature is due to the subcritical speed of the bicycle, which loses further speed with each oscillation.

Figure 13: An example of a dataset, which could be traditionally analysed [27]

4.1 Introduction

"Analytical techniques" is an ill-defined term. For the purposes of this essay we will tighten its meaning to that of techniques which are problem-specific, following analysis of the problem, as opposed to machine learning techniques, which can be applied to many different types of problem without much modification.

Since the scope of analytical techniques is so broad, we will only consider a few specific examples, including mathematical analysis, genetic algorithms, and tree search algorithms - all of which can be used to play games.

4.2 Example of mathematical analysis - Notakto

Not all games can be mathematically analysed, in fact, it is the minority that can be analysed yielding helpful results. However, when it is possible, they produce provably optimal algorithms which use little processing power, which is preferable over many other computational techniques, at the cost of human time spent doing the analysis.

A commonly used example is the game Nim [28] [29], which can be played optimally using a technique known as "nimbers". However, it is interesting to consider less well-known analyses for the sake of variety. Instead, we will consider the game Notakto, which is similar to traditional tic-tac-toe, except with both players placing the same counter (*impartial* play), and the player who makes three in a row loses (*misère* play).

It was initially mentioned in a Math Overflow thread [30], and then solved by Plambeck and Whitehead [31].

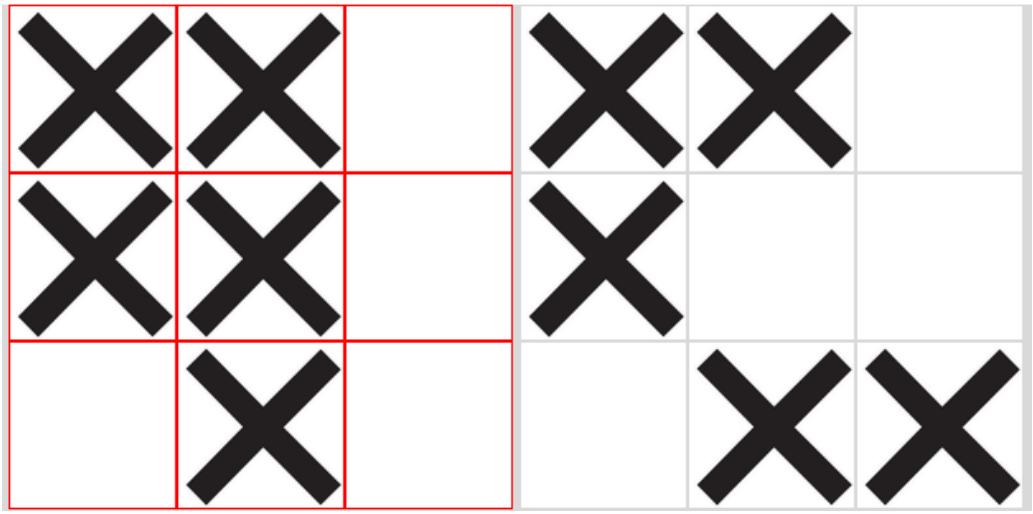


Figure 14: Two Notakto boards, with the left having been won, and the right still in play

It is trivial to find the optimum strategy for a game of Notakto involving only one 3x3 board after playing a couple of games by inspection of two starting cases:

1) If the opponent player plays first with any starting move other than the centre, the player can play the move mirrored through a centre line, and continue doing so until they win.

2) If the player plays first, it should always play in the centre position of the board, then a "knight's move" away from the opponent's piece, and continue doing so until they win.

However, if play is extended to 3 concurrent boards, the problem gets much more complicated. A mathematical technique has been developed to play this game, from a more general approach to other *misère* games.

In short, Plambeck and Whitehead suggest that a certain commutative monoid Q exists that is the *misère quotient* of the game.

$$Q = \langle a, b, c, d \mid a^2 = 1,\ b^3 = b,\ b^2c = c,\ c^3 = ac^2,\ b^2d = d, cd = ad,\ d^2 = c^2 \rangle$$
$$Q = \{1, a, b, ab, ab^2, c, ac, bc, abc, c^2, ac^2, bc^2, abc^2, d, ad, bd, abd\}$$

They then provide a lookup table for all 102 possible *non-isomorphic* board positions with elements in the monoid Q.

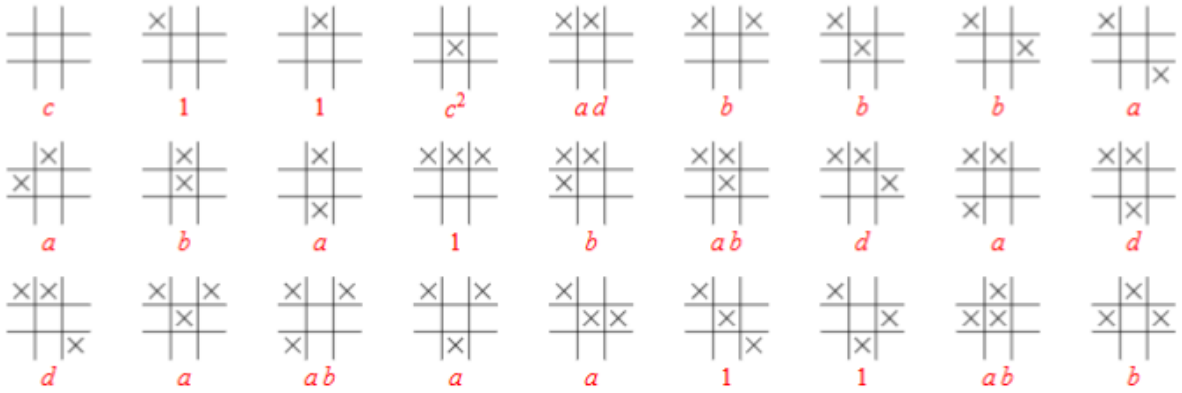


Figure 15: The first 27 items in the lookup table

They then posit that the outcome of the board can be determined by checking if the product of the three elements is equivalent to an element in the set P.

$$P = \{a, b^2, bc, c^2\}$$

This can be used to determine the best move by considering all possible next moves, and selecting one which is predicted to win.

I implemented a computer system to model this, which is included in appendix 7.2 (page 26), and shows how effective and fast this technique is.

This approach allows a player to always win, however, it was very time-consuming to derive, and is only applicable to a single game. Furthermore, it requires the memorisation of 102 values for game boards, which is impractical for most players. However, it does show that abstract games can be approached and solved mathematically, and whilst the solution may be unwieldy, it can be guaranteed to be correct, unlike other algorithms.

4.3 Example of genetic algorithms - Gardner hexapawn

Another approach to play games optimally is the use of genetic algorithms. Genetic algorithms are designed to optimise a model, based on the way that biological systems do, namely genetics and evolution.

These techniques tend to be closer to neural networks that mathematical techniques, in that they are trained, rather than being the solution to the problem immediately. As a result, they can be quite processor and memory intensive, especially on games with a large search space.

We will consider a simple example which was proposed by the popular mathematician Martin Gardner in his column in the Scientific American [32], as applied to a simple game which he called “Hexapawn”:

I have designed hexapawn, a much simpler game that requires only twenty-four boxes. The game is easily analyzed—indeed, it is trivial—but the reader is urged *not* to analyze it.

Hexapawn is played on a 3 × 3 board, with three chess pawns on each side as shown in Figure 43. Dimes and pennies can be used instead of actual chess pieces. Only two types of move are allowed: (1) A pawn may advance straight forward one square to an empty square; (2) a pawn may capture an enemy pawn by moving one square diagonally, left or right, to a square occupied by the enemy. The captured piece is removed from the board. These are the same as pawn moves in chess, except that no double move, *en passant* capture or promotion of pawns is permitted.

The game is won in any of three ways:

1. By advancing a pawn to the third row.
2. By capturing all enemy pieces.
3. By achieving a position in which the enemy cannot move.

Players alternate moves, moving one piece at a time. A draw clearly is impossible, but it is not immediately apparent whether the first or second player has the advantage.

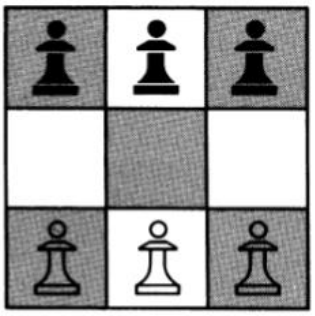


Figure 43
The game of hexapawn

Figure 16: Gardner's summary of hexapawn [33]

The purpose of this game is to have a small search space, in order to make it easier to physically manufacture the genetic algorithm. Gardner then describes the procedure to make and protocol to operate a “HER - hexapawn educatable robot”, using only matchboxes and beads, as contemporary readers didn’t have personal computers.

The basic principle is that moves are randomly selected, and then weighted positively or negatively according to winning or losing the game. After many iterations, all the moves that could lead to a loss are removed, and hence the system can only win.

This algorithm works well on games with small search spaces, such as hexapawn, especially when they can be cut down even further by removing symmetric duplicates. However, on complicated games with large search spaces they don’t work well, as training essentially involves enumerating the entire game tree, which may not be possible.

4.3.1 Differing results

I made an implementation of the genetic algorithm following Gardner’s algorithm, and found that it could train itself to play optimally. The full source code, and description of the algorithm is available in appendix 7.3 (page 28). However, it takes many more iterations than Gardner suggested.

Gardner suggested that “HER” trained to optimal play in 36 iterations, with 11 losses. I found that my simulation trained in 820 iterations, with 62 losses.

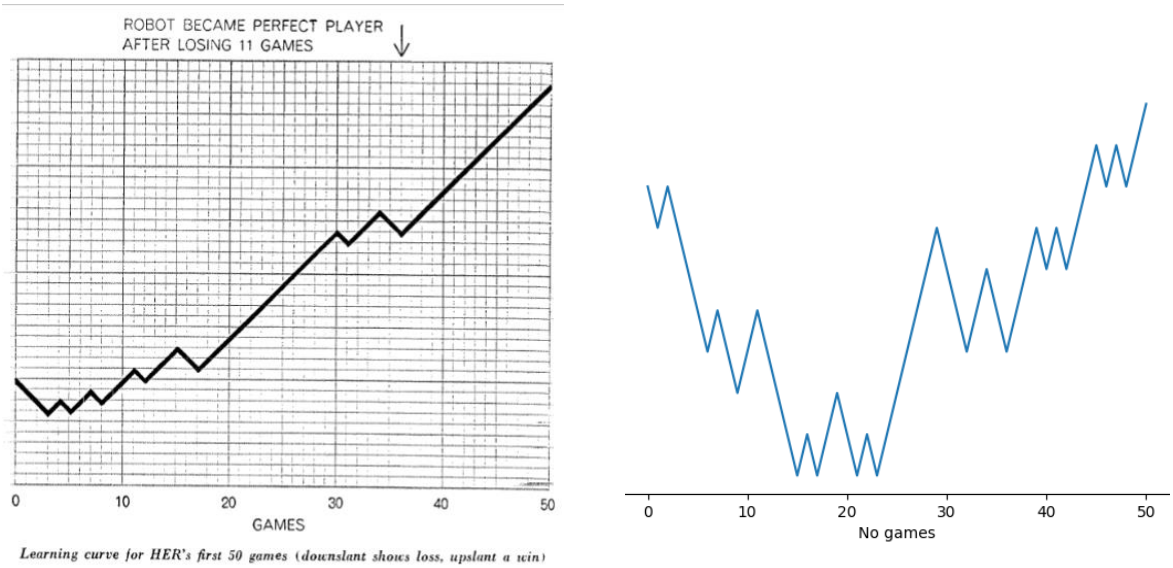


Figure 17: Plots of Gardner’s and our results for the first iterations during training, which look very different

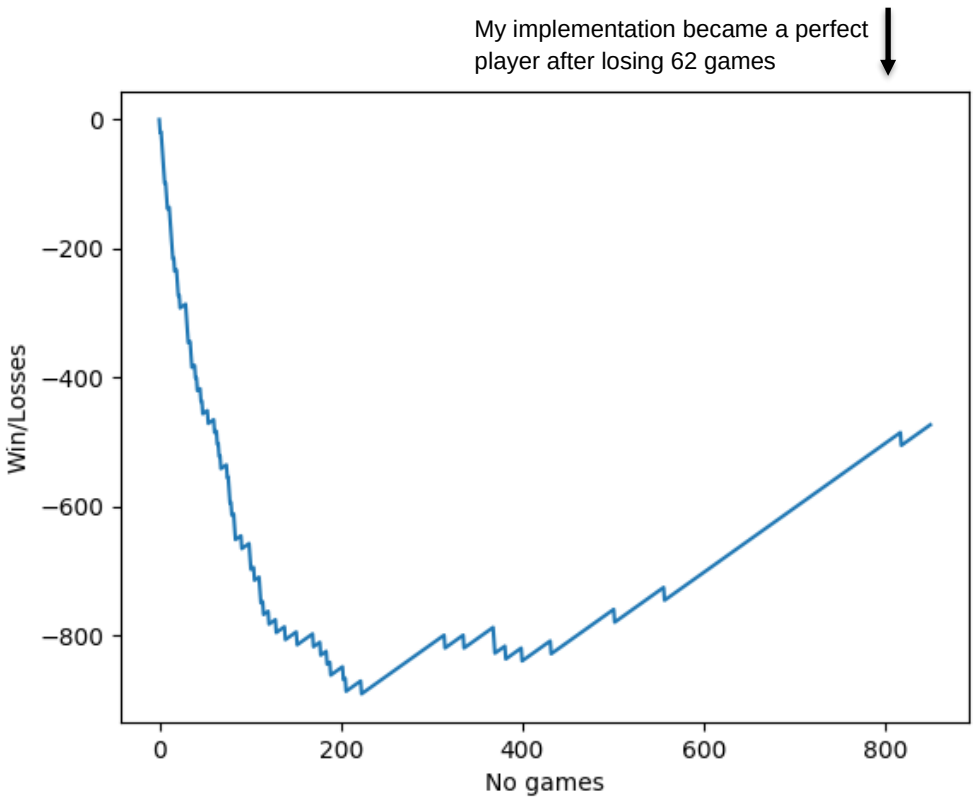


Figure 18: The training curve, with losses being scaled to be more prominent (1:20 ratio)

An important issue to note is Gardner labelled his matchboxes using only *non-isomorphic* boards, as before with Notakto. Initially I didn’t model this, and found it took about 1600 iterations, with 123 losses, to train. I then changed my data structure, reaching the aforementioned 820 iterations with 62 losses.

It is not stated whether play was against a human, and for ease of training I used a random algorithm to pick moves. We suggest that a random training algorithm ended up exploring every possible game state, whereas the tournament training might have only explored a few, since the human opponent would tend to play “sensible” moves.

It is possible that Gardner’s statement the system became a perfect player is erroneous, as he was training manually, which is very time consuming, and he only played 14 games after the supposed turning point. This is indicated by the fact that my model lost 62 games before perfection, so either my implementation is different to his description, or he made an erroneous conclusion, in that some moves can still result in a loss, so it is not a perfect player.

4.4 Example of tree search algorithms – Tic-tac-toe

One of the most commonly used techniques for solving games is by treating them as a tree of game states, where each state is a node, with nodes that are possible next game states after a move:

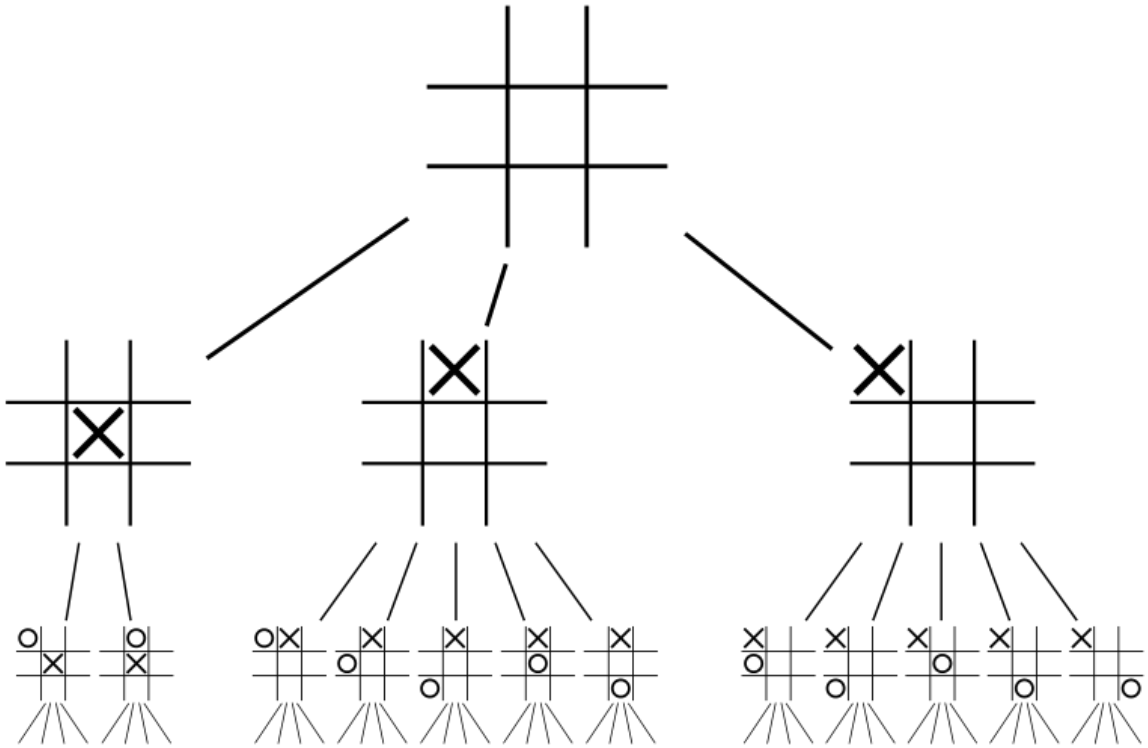


Figure 19: The first three layers of a tic-tac-toe game tree [34]

Since tree algorithms are very fundamental and well understood [35], they are a convenient data structure to apply various algorithms to find optimal solutions.

The simplest way to do this is called the minimax algorithm. The minimax algorithm seeks to pick a decision that **mini**mises the number of ways the opposing player can win, and **max**imises the number of ways the current player can win.

It does this by recursively traversing the game tree. For odd-numbered layers, it tries to minimise the ways to win, and for even-numbered layers, it tries to maximise. It does this by assigning nodes values based on the sum of the values leaf nodes they lead to, with a positive value for a win, and a negative value for a loss. This summation propagates upwards, and the node with the highest value is the node which is most likely to lead to the current player winning.

The pseudocode for the depth limited minimax algorithm is given below.

```
function minimax(node, depth, maximizingPlayer) is
    if depth = 0 or node is a terminal node then
        return the heuristic value of node
    if maximizingPlayer then
        value := −∞
        for each child of node do
            value := max(value, minimax(child, depth − 1, FALSE))
        return value
    else (* minimizing player *)
        value := +∞
        for each child of node do
            value := min(value, minimax(child, depth − 1, TRUE))
        return value
```

```
(* Initial call *)
minimax(origin, depth, TRUE)
```

Figure 20: The formal pseudocode describing the minimax algorithm [36]

I implemented the minimax algorithm to play tic-tac-toe, which I detail in appendix 7.4 (page 30), and found it was very effective, beating some of my peers, and ensuring a draw. The algorithm explores every branch of the game tree, and hence is guaranteed to yield the optimal move. However, for more complex games, it can take a very long time as the game tree gets very big. There are various optimisation techniques to make it faster, which I also discuss in appendix 7.4.1 (page 32).

In summary, tree search algorithms work very well applied to abstract games, and generally the most commonly used in high performance systems, as discussed in the next section.

5 Instances of the application of algorithms to games

Alongside the technical aspects of machine learning algorithms and analytical techniques, I researched cases in which they have already been applied to playing abstract games.

5.1 Applications of analytical techniques

Chinook is a computer system designed to play checkers, developed between 1989 and 2007, and is the first computer program to win a human world championship [47]. It was entirely based on human knowledge, rather than using any machine learning. Chinook first uses an “opening book”, with moves taken from games played by grandmasters. Next, it uses a deep tree search algorithm, which has a board evaluation function using heuristics chosen by humans. Finally, there is a similarly constructed end-game database, for all boards with fewer than eight pieces. This results in a system that is a mix of mathematical analysis, and tree search algorithms – making it a useful example as it is simple and all the heuristics are hand-written, and hence understandable.

Initially, the aim for the project was to beat the best human players. It drew against “the best checkers player of all time” Marion Tinsley [48], and following his withdrawal due to pancreatic cancer, it beat Don Lafferty 1-0. Following this, the projects aim was shifted to solving checkers, which was finally completed in September 2007 [49]

There are many other systems which apply analytical techniques to games. One of the most notable is Stockfish, which competed against AlphaZero, a machine learning approach, discussed sections 5.2 and 5.3 (pages 21 and 22).

5.2 Applications of machine learning techniques

AlphaGo was a project to produce a system that could beat the best human Go players in the world, by Google’s DeepMind [37] [38]. Go was chosen, as it is viewed as one of the most complex board games played by humans, and no other computer system had beaten a human professional on a full-sized board without handicaps.
In March 2016, it played a well-publicised game against Lee Sedol, a 9th dan (the highest rank) Go player, where it won 4-1. In the 37th move of game two, it made a move that “no human ever would”, which left many astounded. One commentator exclaimed, "That's a very strange move", but the system proceeded to win the match. It was later calculated there was a “one in ten thousand” chance a human would have made that move. This highlights one of the ways neural networks can be incredibly effective, they can produce creative, innovative, outputs strongly contrasting analytical techniques, which will not. However, in game four, Lee Sedol also played an equally surprising move, which AlphaGo did not expect any human to make, and was hence

unprepared for it, and lost the game. In summary, we can see that at the very cutting edge of the use of neural networks to play games, both human and computer can produce creative “beautiful” moves, however, modern computers win almost all the time [39].

Following AlphaGo, there were three successor projects, AlphaGo Master, AlphaGo Zero, and AlphaZero. The later versions both took “Zero knowledge” approaches, where there was no human interaction in the training, and they both only used neural networks, as opposed to earlier systems, which used analytical techniques in tandem. These later systems show that neural networks on their own can be incredibly powerful tools, which, in some cases, can exceed even the best analytical techniques.

5.3 Competing machine learning with analytical techniques

AlphaZero was shown to be immensely powerful after beating the foremost chess playing computer system Stockfish in a 100 round face off, winning 25 games as white, 3 as black and drawing the remaining 72.

In that tournament Stockfish used *MCTS* to evaluate 70 million board states, which contrasts AlphaZero only evaluating 80 thousand [40]. AlphaZero could do this since it used machine learning techniques to filter down the set of boards to explore.

This idea of filtering board states, and only exploring promising ones is discussed on page 286 of *Gödel Escher Bach*, where Hofstadter states that “chess masters perceive the distribution of chunks […] a higher level description of the board than [the position of individual pieces]”, following research that showed masters could remember board states faster, and that “masters’ mistakes involved groups of pieces […] which left the game strategically the same”, and confirmed this by the fact that “masters were […] no better than novices in reconstructing random boards.” He suggests that only considering high-level strategically valid moves, rather than just legal ones by master players in a process of “implicit pruning” is how masters play so effectively. [41] This book from 1979 seems to describe the process AlphaZero uses to select moves with remarkable foresight, and shows the progress of machine learning from formulation to realisation.

Many commentators and chess authorities were impressed by the achievement, with Garry Kasparov, who was previously defeated by Deep Blue (the first computer to win at chess against a reigning champion) saying “it's a remarkable achievement, even if we should have expected it after AlphaGo” [42] and the Danish grandmaster Peter Heine Nielsen likened it’s play to that of a superior alien species [43].

However, there are criticisms of the game between Stockfish and AlphaZero, as AlphaZero trained for 9 hours before the match on specialised hardware called TPUs

[44], whereas Stockfish only ran in the match, and it ran on hardware that was stated to be sub-optimal for the task [45]. Finally, others in the computer chess community stated that a newer version of Stockfish was likely to beat AlphaZero [46]. Nonetheless, the game acted as a very important litmus test for both camps.

6 Conclusion

In my research I sought to consider: “How effective are machine learning algorithms compared with traditional analytical techniques, with respect to playing abstract games?”

I found that both machine learning algorithms and traditional analytical techniques can be very effective at playing abstract games. The most modern approaches to complex games tend to fall into two schools of thought, one employing advanced tree search algorithms, often aided by machine learnt heuristics, and the other exclusively employing neural networks. Currently, we do not know which approach is stronger, however, we can be reasonably sure that it will include machine learning algorithms to an extent, if not totally.

In my experience of writing software to demonstrate and better understand this, I found that for very simple games, traditional analytical techniques tend to be faster, and will find the optimal strategy. However, for very complex games, where it is not possible to exhaustively search the entire game tree, nor mathematically analyse play, the more lenient, heuristic nature of neural networks tends to work better.

In summary, machine learning techniques are being used increasingly to play abstract games over or in combination with traditional analytical techniques, with very promising results. This might be in part as they are popular for researchers, as they are able to secure funding, however, they have also proven to be very effective in their own right, and have far exceeded human level of play.

7 Appendices

In the appendices, I predominantly address the technical aspects of software I wrote to better understand my topic of research. All the code I have referenced throughout this essay is written in python, predominantly conforms to PEP8 [50], and is available on a public GIT server: https://github.com/EdmundGoodman/rouse-research

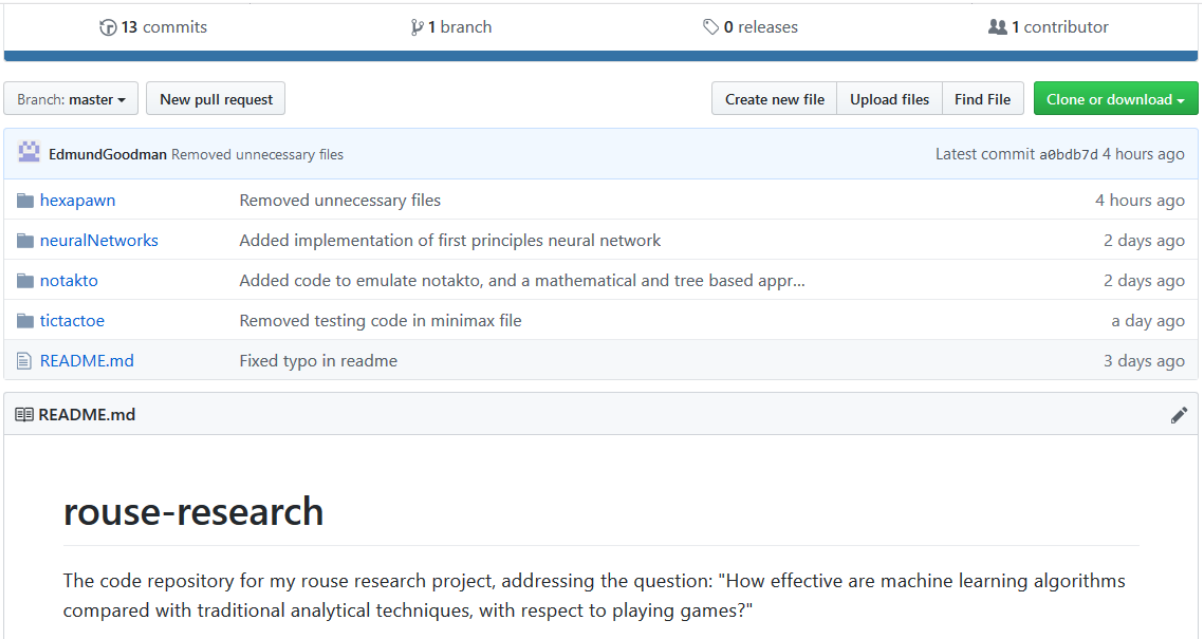


Figure 21: A screenshot of the git server webpage

7.1 Training first principles neural networks

In order to ensure that our implementation of the neural network functioned as expected, and to get a better understanding of the process of training before attempting to play tic-tac-toe, we trained the network on a common benchmarking dataset: IRIS [51]

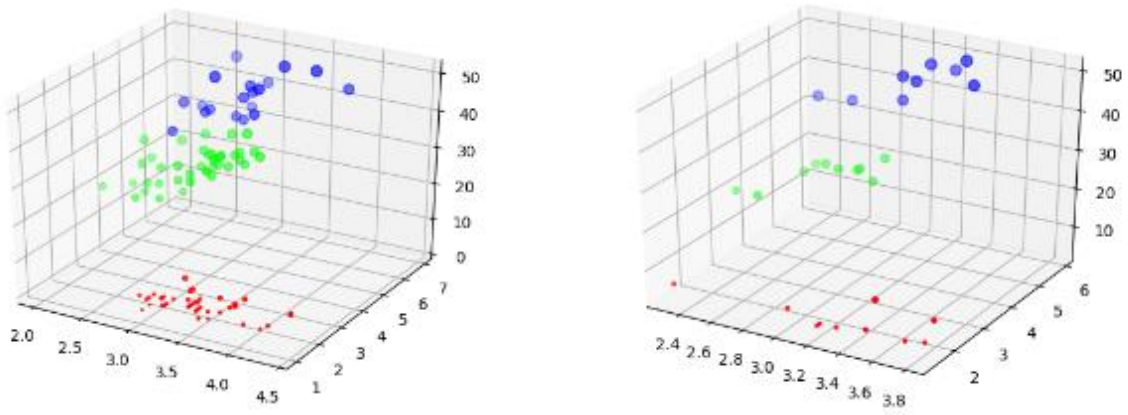


Figure 22: A graph of the training (left) and testing (right) data in the iris data set

The IRIS data set is a multivariate set introduced by Ronald Fisher in 1936, which is now used as a simple classification problem for testing neural networks. The expected categories are given by colour in the above graph, and the data encoded in 3-axis and size.

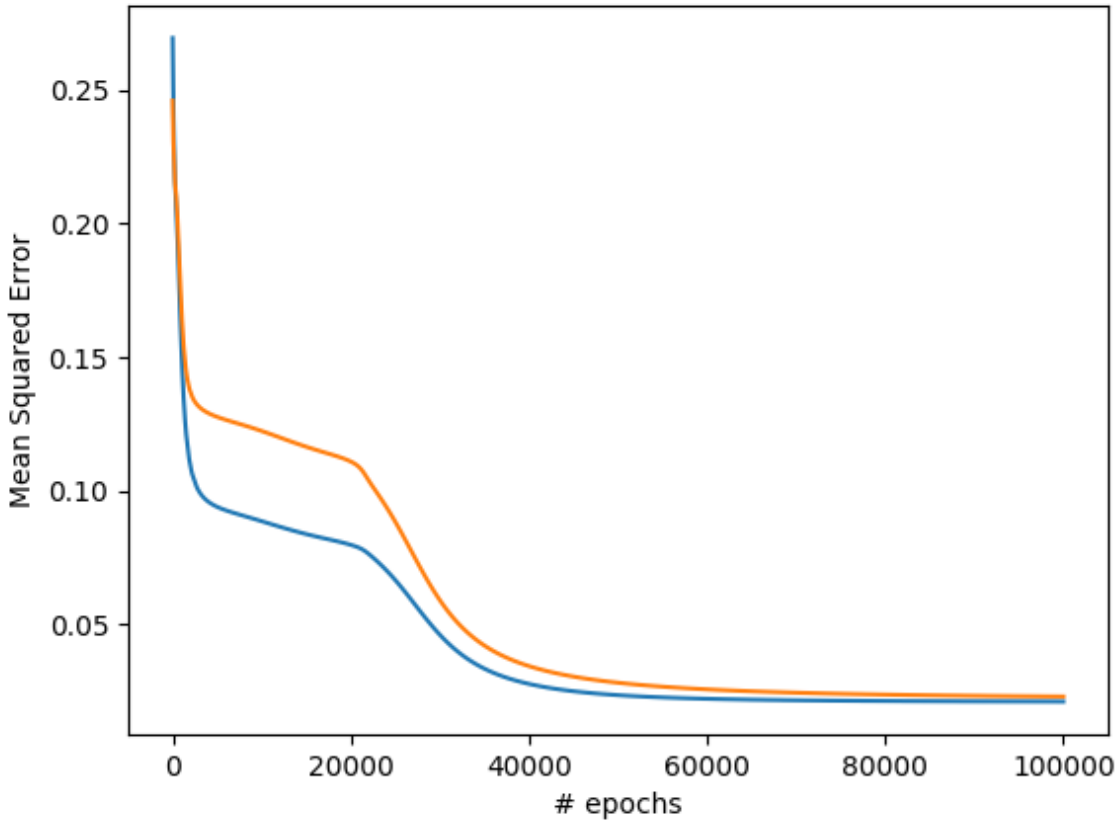


Figure 23: A graph of the mean squared error during training, with the blue line being testing, and the orange training data

We calculated mean squared error during training, and plotted it on a graph, which can be seen to converge to a very small value, indicating successful training. We yielded a 98.0% training accuracy and a 100.0% testing accuracy – confirming our initial success.

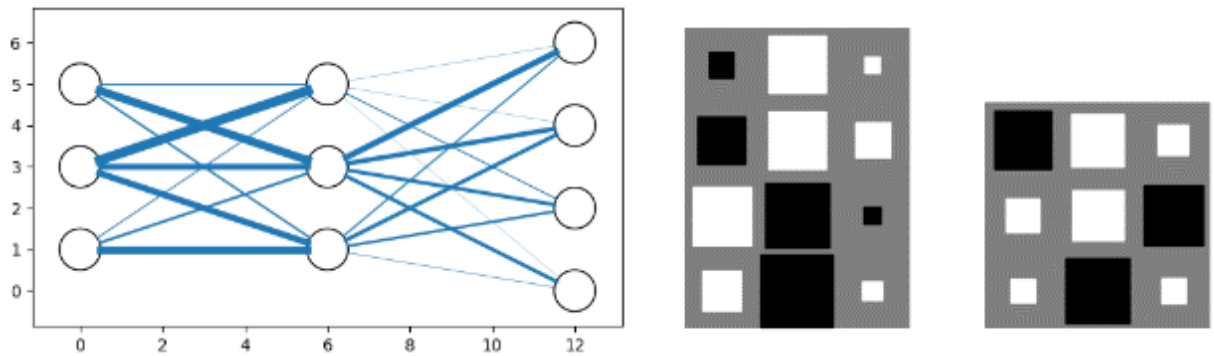


Figure 24: A representation of the trained network as a network diagram (left) and Hinton diagram (right)

We then wrote software to visualise the network as network and Hinton diagrams, so we could see the internal workings of the system, and verify it works.

7.2 Implementing mathematical analysis to solve Notakto

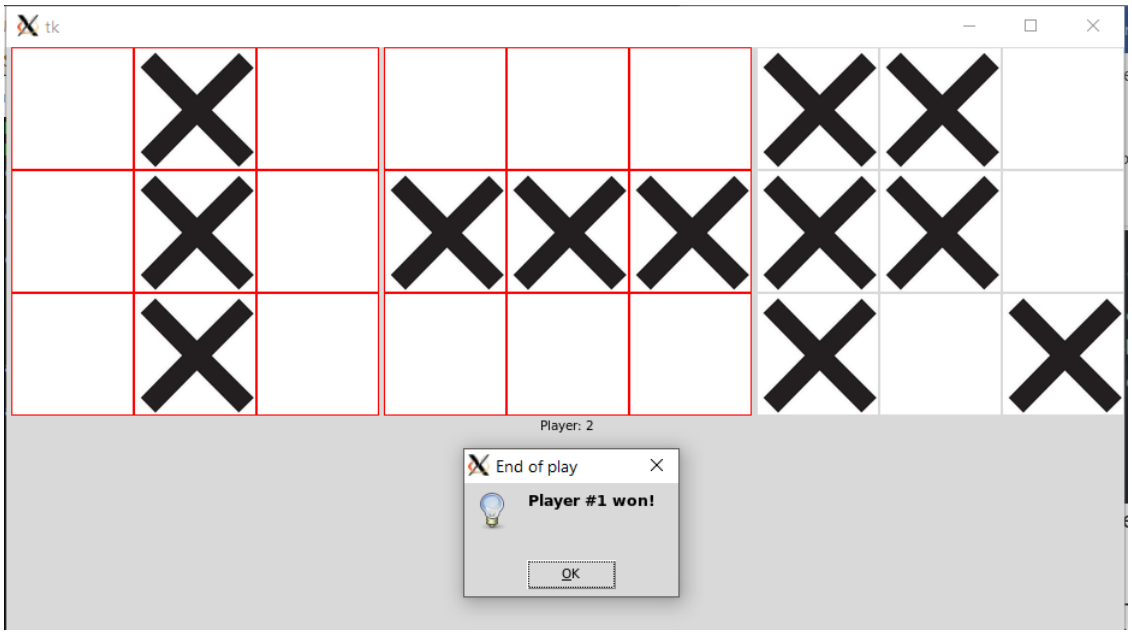
I implemented a computer program to graphically model 3 board, 3x3 Notakto, which allowed both human vs human, and more interestingly human vs AI play. Since the approach relied on the paper involves memorising a vast set of board states, and their corresponding polynomial terms I thought that it would be a good fit for a computer implementation, as the main problem humans face with the memorising this data is easy for a computer. I manually translated all the board states to a python hash table of each distinct board state and its complementary polynomial, along with the transformations in the definition of the monoid Q, and the target safe states as a list.

```
data = {
    ((0,0,0),(0,0,0),(0,0,0)):"c",
    ((0,0,0),(0,1,0),(0,0,0)):"cc",
    ((1,1,0),(0,0,0),(0,0,0)):"ad",
    ((1,0,1),(0,0,0),(0,0,0)):"b",
    ((1,0,0),(0,1,0),(0,0,0)):"b",
```

```
reductions = {"aa":"","bbb":"b","bbc":"c",
              "ccc":"acc","bbd":"d","cd":"ad",
              "dd":"cc","":"aa", "b":"bbb",
              "c":"bbc","acc":"ccc","d":"bbd",
              "ad":"cd","cc":"dd"
}

targets = "a bb bc cc".split()
```

First, I produced a GUI for the game and tested it in human play, as this feature would speed up debugging later on, then moved onto the AI component.



I realised that the process of mutating the polynomials to see if they were in the set of safe states was as non-trivial for a computer as a human.

In order to check if the board is a safe state, I concatenated the polynomial states of each board, then implemented a depth limited tree search of the possible

transformations, until either a target state or the recursion depth is hit, indicating whether or not the string is able to be reduced to a target state.

```
265     def getValidChildren(self, node):
266         children = []
267         #Apply one substitution at a time, and defer nested over the tree
268         for string,substitution in reductions.items():
269             start = node
270             for char in node:
271                 oldString = string
272                 string = string.replace(char,"",1)
273                 if oldString != string:
274                     start = start.replace(char,"",1)
275             if string == "":
276                 children.append("".join(sorted(start+substitution)))
277             else:
278                 pass
279         return list(set(children))
```

```
281     def reducePolynomial(self, polynomial):
282         polynomial = "".join(sorted(polynomial))
283         count, tree, visited = 0, [polynomial], [polynomial]
284         while True:
285             #print(count, tree)
286             next = []
287             for node in tree:
288                 if node in targets:
289                     return node
290                 next.extend(self.getValidChildren(node))
291             next = [n for n in next if n not in visited]
292             visited.extend(next)
293             tree = next[:]
294             #If the recursion depth is exceeded
295             if count > 3 or len(tree) == 0:
296                 return sorted(visited, key=lambda x: len(x))[0]
297             count += 1
```

Using the function to check if the board state was safe, the AI then randomly selects moves from the safe board states, along with a heuristic suggested by the author for human play of “killing” boards if possible.

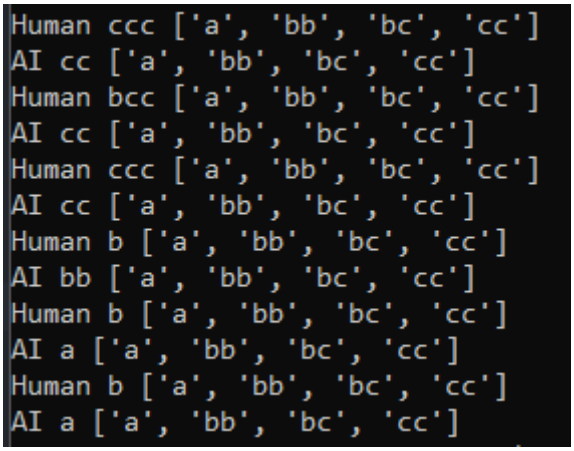


Figure 25: The AI player maintaining a safe board state throughout a game

I found that the finished product exceeded my expectations in terms of how effectively it played, however I experienced various bugs during testing, for example errors in the data file, and conceptual errors in the algorithm. I have tested the finished product to an extent, but I cannot easily verify if it is perfect, due to the massive search space, and as I previously found bugs it is important to consider the possibility there may be more lying undiscovered.

7.3 Implementing genetic algorithms to solve Hexapawn

Gardner initially suggested making a physical system to act as a genetic algorithm, however, making the physical system is time consuming, and playing many games against it is more so. As a result of this, I implemented a software version of the system, with an object modelling the matchbox computer.

```
322     def getComputerMove(self, board, playerNum, count):
323         boardTuple = tuple([tuple(elem) for elem in board.getBoard()])
324         mirrorTuple = self.getMirroredBoard(boardTuple)
325 
326         mirrored = mirrorTuple in self.boardStates
327 
328         if (boardTuple not in self.boardStates) and (mirrorTuple not in self.boardStates):
329             #If we haven't encountered this move before
330             self._openBoxes.append(boardTuple)
331             moves = board.getValidMoves(playerNum)
332             defaultWeight = {1:4, 3:3, 5:2, 7:1}[count]
333             movesWeights = {move:defaultWeight for move in moves}
334             self.boardStates[boardTuple] = movesWeights
335         else:
336             #If it is a normal board
337             if mirrored:
338                 boardTuple = mirrorTuple
339 
340             self._openBoxes.append(boardTuple)
341             movesWeights = self.boardStates[boardTuple]
342             moves = []
343             for key,value in movesWeights.items():
344                 moves.extend([key]*value)
```

```
346         if len(moves) == 0:
347             self._resigned = True
348             return "Resign", "Resign"
349 
350         move = random.choice(moves)
351         self._movesMade.append(move)
352         if mirrored:
353             return self.getMirroredMove(move)
354         else:
355             return move
356 
```

As shown above, I modelled the matchboxes as a hash table with keys containing the tuple representation of the board, and the value being another hash table with keys containing the tuple representation of the move, and the value its weight.

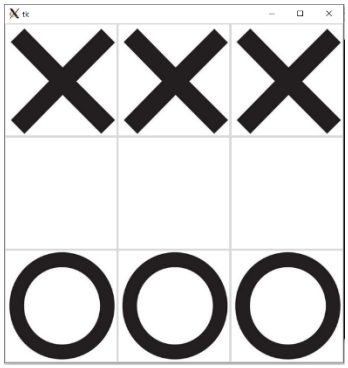
```
313     def getMirroredBoard(self, board):
314         newBoard = []
315         for row in board:
316             newBoard.append(tuple(reversed(row)))
317         return tuple(newBoard)
318 
319     def getMirroredMove(self, move):
320         return ((2-move[0][0], move[0][1]),(2-move[1][0], move[1][1]))
321 
```

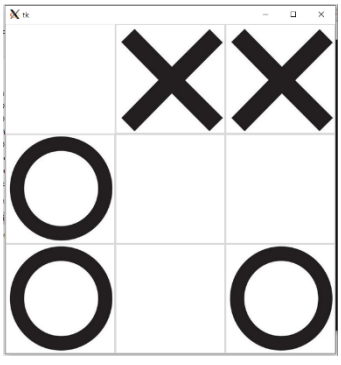
I also check and removed isomorphically duplicate boards, as discussed before, to more closely fit the physical model suggested by Gardner

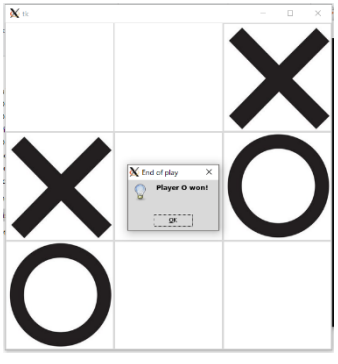
```
296     def updateWeights(self, won):
297         if not won:
298             index = -2 if self._resigned else -1
299             self.boardStates[self._openBoxes[index]][self._movesMade[-1]] += -1
300         self.resetGame()
301         self.cleanUpBoardStates()
302 
303     def cleanUpBoardStates(self):
304         newBoardStates = {}
305         for Mkey in self.boardStates.keys():
306             newMovesWeights = {}
307             for key, value in self.boardStates[Mkey].items():
308                 if value > 0:
309                     newMovesWeights[key] = value
310             newBoardStates[Mkey] = newMovesWeights
311         self.boardStates = newBoardStates
312 
```

During training, I maintained the matchbox computer by changing weights as Gardner described following a loss.

Again, I implemented a GUI to play test the AI against, as shown below, with the AI playing second and winning



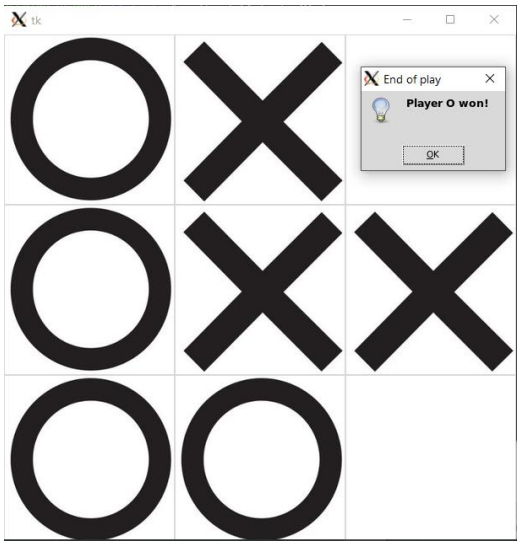




7.4 Using tree search algorithms to play tic-tac-toe

I attempted to fully understand the simplest version of the minimax algorithm by implementing it to play tic-tac-toe, and hence it is a good direct comparison to similar neural network solution discussed later to the same problem.

Firstly, implemented a GUI for tic-tac-toe:



And following this, implemented the minimax algorithm passing the board state up a recursive call stack as a method of my game class:

```
170     def minimax(self, board, players, utilities=[1,-1,0,0], leafDepth=9, depth=0):
171         if depth>leafDepth: #Simulate leaf nodes beyond recursion depth
172             return utilities[3]
173 
174         winner = board.checkIfWon()
175         if winner is not False:
176             if winner == players[0]: #Win
177                 return utilities[0]
178             elif winner == players[1]: #Loss
179                 return utilities[1]
180             else: #Draw
181                 return utilities[2]
182 
183         moveWeights = {}
184         for move in board.getEmptySquarePositions():
185             nextBoard = TictactoeBoard()
186             nextBoard.setBoard(board.getBoard())
187             nextBoard.setPlayerNum(board.getPlayerNum())
188             nextBoard.makeMove(move)
189             nextBoard.togglePlayer()
190 
191             moveWeights[move] = self.minimax(nextBoard, players, utilities, leafDepth, depth+1)
```

```
193         if depth%2==0: #If it is the ai playing, play best move for the ai
194             bestMove = max(moveWeights, key=lambda key: moveWeights[key])
195         else: #Otherwise, play the best move the other player can make (worst move for the ai)
196             bestMove = min(moveWeights, key=lambda key: moveWeights[key])
197         bestWeight = moveWeights[bestMove]
198 
199         if depth == 0:
200             return bestMove
201         else:
202             if True:
203                 return bestWeight
204             else: #If playing against non-optimal opponent, consider:
205                 return sum(moveWeights.values())
```

Furthermore, I implemented a helper function to yield the optimal move given the current board state

```
207     def getBestMove(self):
208         duplicateBoard = TictactoeBoard()
209         duplicateBoard.setBoard(self.board.getBoard())
210         duplicateBoard.setPlayerNum(self.board.getPlayerNum())
211         x = time.time()
212         val = self.minimax(duplicateBoard, duplicateBoard.getPlayers())
213         print(val, round(time.time()-x, 5))
214         return val
```

I then tested the algorithm thoroughly, and found that it worked incredibly effectively, and never lost against me or any other tester. However, it takes a long time (about 5 seconds) to evaluate the first move.

7.4.1 Further tree search algorithm techniques

The minimax algorithm can be improved by employing alpha-beta pruning, which is a heuristic which reduces search time by searching more promising paths first.

The pseudo-code for minimax with alpha-beta pruning is as follows:[12]

```
function alphabeta(node, depth, α, β, maximizingPlayer) is
    if depth = 0 or node is a terminal node then
        return the heuristic value of node
    if maximizingPlayer then
        value := −∞
        for each child of node do
            value := max(value, alphabeta(child, depth − 1, α, β, FALSE))
            α := max(α, value)
            if α ≥ β then
                break (* β cut-off *)
        return value
    else
        value := +∞
        for each child of node do
            value := min(value, alphabeta(child, depth − 1, α, β, TRUE))
            β := min(β, value)
            if α ≥ β then
                break (* α cut-off *)
        return value
```

```
(* Initial call *)
alphabeta(origin, depth, −∞, +∞, TRUE)
```

Figure 26: The pseudocode for minimax with alpha-beta pruning [52]

I also updated my minimax code to increase speed using alpha-beta pruning:

```
170     def minimax(self, board, players, utilities=[1,-1,0,0], alpha=-math.inf, beta=math.inf, leafDepth=9, depth=0):
171
172         if depth>leafDepth: #Simulate leaf nodes beyond recursion depth
173             return utilities[3]
174
175         if depth==0:
176             maxValue, bestMove = -math.inf, None
177             for move in board.getEmptySquarePositions():
178                 nextBoard = TictactoeBoard()
179                 nextBoard.setBoard(board.getBoard())
180                 nextBoard.setPlayerNum(board.getPlayerNum())
181                 nextBoard.makeMove(move)
182                 nextBoard.togglePlayer()
183
184                 value = self.minimax(nextBoard, players, utilities, alpha, beta, leafDepth, depth+1)
185                 if value > maxValue:
186                     bestMove, maxValue = move, value
187
188             return bestMove
189
190
191         winner = board.checkIfWon()
192         if winner is not False:
193             if winner == players[0]: #Win
194                 return utilities[0]
```

```
            elif winner == players[1]: #Loss
                return utilities[1]
            else: #Draw
                return utilities[2]

        if board.getCurrentPlayer() == players[0]: #Maximising layer
            value = -math.inf
            for move in board.getEmptySquarePositions():
                nextBoard = TictactoeBoard()
                nextBoard.setBoard(board.getBoard())
                nextBoard.setPlayerNum(board.getPlayerNum())
                nextBoard.makeMove(move)
                nextBoard.togglePlayer()

                #print(nextBoard)

                func = self.minimax(nextBoard, players, utilities, alpha, beta, leafDepth, depth+1)

                value = max(value, func)
                alpha = max(alpha, value)
                if depth==0: print(move, alpha, beta, 1)
                if alpha >= beta:
                    if depth==0: print("Hit 1")
                    break

```

```
        else:
            value = math.inf
            for move in board.getEmptySquarePositions():
                nextBoard = TictactoeBoard()
                nextBoard.setBoard(board.getBoard())
                nextBoard.setPlayerNum(board.getPlayerNum())
                nextBoard.makeMove(move)
                nextBoard.togglePlayer()

                func = self.minimax(nextBoard, players, utilities, alpha, beta, leafDepth, depth+1)

                value = min(value, func)
                beta = min(beta, value)
                #if depth==1: print(move, alpha, beta, 2)
                if alpha >= beta:
                    #if depth==1: print("Hit 2")
                    break

        if depth == 0:
            return move
        else:
            return value

```

And again added a helper function to yield the best move for a given board position

```
    def getBestMove(self):
        duplicateBoard = TictactoeBoard()
        duplicateBoard.setBoard(self.board.getBoard())
        duplicateBoard.setPlayerNum(self.board.getPlayerNum())
        x = time.time()
        val = self.minimax(duplicateBoard, duplicateBoard.getPlayers())
        print(val, round(time.time()-x, 5))
        #print(val)
        return val

```

It is notable that whilst alpha-beta pruning greatly increases the speed of traversal, it also requires much more source code to be implemented.

Throughout my essay, I mention the Monte Carlo Heuristic Search Tree algorithm (MCTS). This is a more advanced approach, which employs a monte carlo style random algorithm to search the tree, often in combination with uncertainty bounds. This means that it doesn't search every board state, but it searches more promising branches earlier, making it effective for complex games, such as Chess and Go.

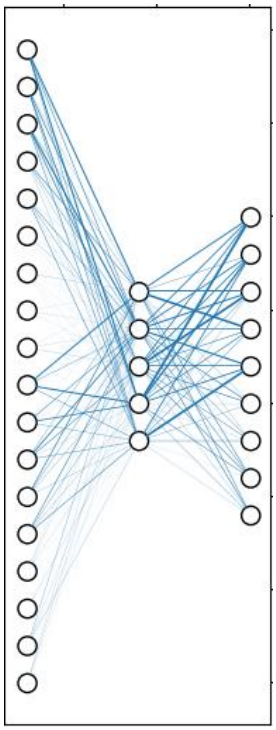
7.5 Using first principles neural networks to play tic-tac-toe

In order to gain a full and broad understanding of neural networks as applied to games, I decided to write a program to try an optimally play a very trivial game: tic-tac-toe. I found rather quickly that writing a neural network to do this was rather difficult. Throughout the process, I exclusively used the first principles network I developed, as detailed above.

I eventually tried two approaches which both worked well: A genetic based algorithm, and a back-propagation algorithm which trained on a set of good moves generated by an optimal player.

The genetic algorithm produced started with a large pool of initial weights, and played them against a random opponent. If the weights lost the round, they were removed from the pool. At a threshold, I randomly "bred" weights together, selecting a small number, then taking the average of them, then added them to the pool. I then repeated these steps until the pool didn't change size 10 consecutive times, i.e. no games had been lost.

The back-propagation algorithm trained off an input and target output set generated from an alpha-beta pruned tree search, which always gave the correct move given the board state. I trained this until it made the correct move close to every time, then that set of weights was my trained network.



It is important to consider what is a good input and output schema to train the network on.

Initially I used a 9 node input layer, with empty squares taking the value zero, “X”s one, and “O”s minus one. However, I found that the sigmoid activation function I was using stripped out the sign of the value, so it couldn’t train.

I then moved to an 18-node input layer, with the first 9 nodes taking the value of one if the square held the AI player, otherwise zero, and another nine nodes doing the same thing the opposing player.

Then, I used a 9-node output layer, with the highest node value indicating where to place the cross worked well.

This larger input data set helped the network avoid underfitting.

I then trained the networks, changing the weights millions of times, in order to find a functional combination.

I also experienced issues in network architecture, and found that small networks could not describe the input set appropriately, and so I used a large 18 x 30 x 30 x 9 network in my final testing.

The full code can be found on GitHub, and both attempts have many optimisations and techniques employed to make them work, but there are too many to explicitly enumerate her.

Both take a long time to train, and owing to the random nature of their training, sometimes end up with flaws, occasionally losing games.

7.6 Further neural network techniques

We have described the basic principles of neural networks, and how they are trained. It is tempting to leave it there, and merely throw more data and processing power at a problem, which it should be noted many researchers have been known to do. However, there are a wide array of techniques which can be used to increase the efficacy of networks, and make them more applicable to different problem sets. It is almost always better to consider these techniques, many of which are not especially complicated.

7.6.1 Overfitting and underfitting

The solution to underfitting is relatively trivial, a different, larger network architecture. However, it can be difficult to identify if a network is underfitting, as the only sign is it not getting better, which could be symptomatic of many different issues, and sometimes it is not feasible to increase the network size, as it means more computation is required per epoch.
The problem of overfitting is that an overfitted network will only be correct for its training data, and hence will not yield a general solution to the problem, as required. It is also more difficult to address, and is already somewhat ill-defined, as there is not a clear line denoting when overfitting starts, and all networks will overfit to some extent, by virtue of the mechanism by which they are trained.

There are an immense number of techniques that can be used to combat overfitting and underfitting, along with other techniques to improve the speed and efficiency of training, and different architectures that have proven to be useful for specific classes of problems. Some of them are enumerated below:

7.6.2 Regularisation

Regularisation is a technique used to mitigate overfitting. It does this by taking note of the fact that if a network overfits, it will end up with large differences in value between individual weights. This means that overfitting can be mitigated by also penalising a network for having very different weights. This is often implemented with a “Lambda” hyperparameter, which denotes how much the network is penalised for weight disparity:

```
89      def updateWeights(self, X, learnRate):
90          #Update the hidden layer weights
91          j = 0
92          for i in reversed(range(len(self.W2))):
93              currDelta = np.dot(self.a[i].T, self.delta[j])
94              currDelta -= (self.Lambda*self.W2[i]) #Regularisation - penalise large weight values
95              currDelta *= learnRate #Learning rate - possible implement momentum
96
97              self.W2[i] += currDelta
98              j += 1
99
100         #Update the input layer weights
101         currDelta = np.dot(X.T, self.delta[j])
102         currDelta -= (self.Lambda*self.W1) #Regularisation - penalise large weight values
103         currDelta *= learnRate #Learning rate  - possible implement momentum
104
105         self.W1 += currDelta
```

7.6.3 Hinton’s dropout

A second solution to this is called Hinton’s dropout [53], and is remarkably simple in its function, and is a relatively new discovery. During the forward propagation step of the training process, a small proportion of hidden nodes are randomly turned off. This

effectively introduces random noise into the dataset, so irrespective of the number of training passes, the network will not overfit as much. Again, this can be implemented with a metaparameter giving the fraction of hidden nodes that are dropped.

```
53      def doDropout(self, layer, weight):
54          #https://www.cs.toronto.edu/~hinton/absps/JMLRdropout.pdf
55          if weight!=0:
56              layer *= np.random.binomial([np.ones((len(layer),len(layer[0])))],1-weight)[0] * (1.0/(1-weight))
57          return layer
58
```

7.6.4 Pruning

Another technique that can be used to improve the efficiency of training a network is pruning [54]. Pruning is the process of removing the least significant node in a network, in order to increase training speeds, without reducing the effectiveness of a network. If a smaller network can achieve the same accuracy of output of a larger network, it should be preferred, and pruning provides a way to reduce the size of a network architecture, without dropping effectiveness

7.6.5 Stochastic & Batch processing

There are two main modes of learning for neural networks, stochastic and batch processing. In stochastic learning, the weights are updated after each training example, based on only that error, whereas in batch learning the errors are accumulated over a set of training examples, and the weights are updated at the end. Stochastic learning introduces “noise” into the descent, with each individual update going in slightly different directions, whereas batch learning tends to result in a less “noisy” descent, as the change in weights are averaged over the entire batch, however, repeatedly training on the same batch can lead to negative effects. A common technique is to use mini-batches, randomly selected sets of input data from the training set, back-propagated together, which has been shown to provide good training characteristics.

```
#Generate minibatches to train the network with
combinedData = [(X[i].tolist(), y[i].tolist()) for i in range(len(X)-1)]
random.shuffle(combinedData)
batchX = np.array([d[0] for d in combinedData[:self.batchSize]])
batchY = np.array([d[1] for d in combinedData[:self.batchSize]])

yHat = self.nn.forwardpropagate(batchX)
self.nn.backwardpropagate(batchX, batchY, yHat)
self.nn.updateWeights(batchX, self.learnRate)
```

7.6.6 Learning rates

Learning rates are hyperparameters that linearly scale the rate of change of the weights in a network. If the learning rate is too high, the weights change very quickly, and so can "jump" over a minimum, however, if the learning rate is too low, the network will train very slowly. There are techniques which dynamically change the learning rate throughout training to improve the rate of training, such as momentum.

7.6.7 Weight picking

Initialising networks with good weights can influence the training of the network vastly. Weights are generally picked to be small and centered about zero. A common technique for initialisation, called Xavier's initialisation is to generate a normally distributed random sample of weights with mean zero and standard deviation one, then multiply the sample by the square root of the reciprocal of the number of inputs to a given layer [55] [56].

```
return np.round(np.random.randn(x_dim,y_dim),r)*(1/(numIn+numOut))**0.5
```

7.6.8 Activation functions

The logistic sigmoid function discussed in the neural networks section is a very commonly used activation function. However, it does have some drawbacks, including the "vanishing gradients" problem, where in the near horizontal sections of the function, i.e. very large magnitude input values, the gradient in backpropagation is very small, so it trains slowly [57].

There are many activation functions other than the logistic sigmoid curve used in neural networks. Some of the more common ones detailed in "Information theory, inference, and learning algorithms", by David MacKay [19] include:

Linear activation

$$a(x) = x$$

Seems fast, however doesn't function effectively at all. This can be explained in two ways: One is that any sequence of nodes can be expressed as a single node, so deep networks will not be more effective. The second is that the derivative is constant, as $\frac{d}{dx}\big(a(x)\big) = 1$. As a result of this backpropagation, which relies on differentiation, cannot be employed.

Tanh Sigmoid

$$a(x) = \tanh(x) = 2\,sigmoid(x) - 1$$

Similar to sigmoid, but allows negative values

Threshold/Step/Heaviside function

$$a(x) = \Theta(x) = \begin{cases} 1, x > 0 \\ -1, x > 0 \end{cases} = \frac{d}{dx}(\begin{cases} x, x > 0 \\ 0, x > 0 \end{cases})$$

Often used in final layers of binary classifies, as it will yield an output of the correct (boolean) data type

ReLu (Rectified linear unit)

$$a(x) = \begin{cases} x, x > 0 \\ 0, x > 0 \end{cases}$$

The ReLu is non-linear, so it doesn't have some of the problems of a linear activation function, e.g. it can be used with back-propagation. However, it is not bounded, so it can still explode in size. Furthermore, it increases the efficiency of a network being improving a property called sparsity. In other functions, such as sigmoid, every input value yields an analogue output value, whereas ReLu yields a binary 0 about 50% of the time, so it is much less computationally expensive to evaluate it.

However, this binary portion means that the gradient becomes zero, so neurons stop responding to backpropagation. This is called the dying ReLu problem, where some nodes stop responding, and hence reduce the efficiency of the network. To fix this, a "Leaky" ReLu can be employed:

$$a(x) = \begin{cases} x, x > 0 \\ 0.01x, x > 0 \end{cases}$$

Which means neurons don't die. Finally, ReLu can be faster than sigmoid or tanh, since it involves simpler functions – which makes it preferable in the context of some deep networks.

7.6.9 Convolution, deep and recurrent

There are many modern neural network architectures which have been developed to be applied to specific problem sets.

Deep neural networks are an example of this, when many hidden layers are used, which allows for the network to model very complex systems. They are the most common approach to general problems, as they tend to work well, but don't have overly specific optimisations, such as those of CNNs and RNNs. However, they often take a long time to train.

There are also more drastic deviations from the initial basis of neural networks, for example convolutional neural networks, CNNs.

Simple neural networks tend not to work very effectively on large data sets of images. For example, a common test for the effectiveness of a neural network is the MNIST data set, which contains 60,000 28x28 grayscale images of handwritten digits. However, in order to input this into a neural network, a 784 node input layer is required, along with a number of hidden layers, which results in a large network that is very hard to train. Convolutional neural networks help mitigate this problem by having an additional special type of layer bridging between the input layer and the hidden layers, which reduces the amount of processing which needs to be done. It does this by pre-processing the image, and identifying its features, reducing the amount of information going into the main hidden layers. This process has the additional benefit of making the network more resistant to features displaced in the image, which neural networks alone don't handle very well.

Kernel convolution is the process of passing a matrix, known as a kernel over an input signal, in order to transform it into a feature map, dependent on the kernel. The output signal of an individual pixel is the sum of the products of the kernel weight and its corresponding input square, when the kernel is centred on the target pixel.

$$C[x, y] = \sum_i \sum_j A[i, j] \times B[x - i, y - j]$$

Where A is the input signal, B is the kernel, and C is the output signal.

This idea of employing feature maps as a pre-processing layer in deep learning was first published in 1990 by Yann LeCunn "a founding father of convolutional nets" et al. [58], and was then popularised as a viable technique by Krizhevsky, Sutskever and Hinton following the resounding success of the AlexNet entry into the 2010 ImageNet contest [59].

Traditional neural networks cannot handle real time data, as they would not be able to train using backpropagation. However, being able to process real time data would be very useful. The main approach to this problem is to employ recurrent networks, where later layers feed back into earlier ones. This allows the processing of a real time stream of data, and can also simulate memory. This is a complicated issue, which was initially considered intractable, and is best addressed in Ilya Sutskever's seminal thesis [60].

8 Glossary

This area of research includes many pieces of subject specific technical vocabulary, that is either not in normal use, or has a subtly different meaning to its traditional use in the vernacular, as a result of this, below is a glossary of terms to allow improved specificity of language, and hence reduce unnecessary additional wording.

Term	**Definition**
Machine learning	An application of artificial intelligence that allows systems to automatically learn and improve from experience without being explicitly programmed
Analytical technique	A technique produced by human analysis of a problem
Algorithm	A process or set of rules to be followed in calculations or other problem-solving operations
Game	A physical or mental competition conducted according to rules with the participants in direct opposition to each other
Abstract games	A specific type of game, in which the theme is unimportant to the experience of gameplay, and hence can be usefully analysed by computer algorithms.
Neural network	A prominent machine learning technique modelled on the human brain
Regression	A problem involving predicting numerical data given input data, based on its properties in a supervised manner
Classification	A task involving dividing a dataset into classes based on properties of the data in a supervised manner
Clustering	A task involving partitioning a dataset into groups in an unsupervised manner
Underfitting	The problem of a neural network not fitting to a data set closely enough
Overfitting	The problem of a neural network fitting to a dataset so closely it loses applicability to the general problem
Network architecture	The size and structure of a neural network, including number hidden nodes, types of connections, types of activation functions, etc.
Perceptron/Node/Neuron	A single unit used to compose a neural network
Activation function	A non-linear function used in a node in a neural network
Weights	Properties of a neural network that change during training, encoding the model
Meta-parameters	Properties of a neural network which are fixed during runtime, e.g. network architecture
Forward propagation	The process passing data through a network to yield an output
Backpropagation	The process of finding the amount to change the weights by to improve a network
Epochs	A full cycle of forward propagation followed by back propagation
NumPy	A python library for scientific computing, specifically with support for efficient matrix operations
Tensorflow	A symbolic mathematics library that is most commonly used for machine learning, such as neural networks
Tensor processing units	(TPUs), a specialised piece of computer hardware, often used to train neural networks
Misère play	A type of a game where the last person to play a move loses
Impartial play	A type of game where valid moves depend only on position, not which player is moving
Misère quotient	Mathematical structures (commutative monoids) that encode the additive structure of specific misère-play games
Non-isomorphic	Objects that are not structurally equivalent, i.e. are not isomers of each other, for example an impartial game board mirrored in a line of symmetry. (A is isomorphic to B is denoted by $A \cong B$)
Game tree	A directed graph, whose nodes are game states, and whose edges are moves
Search space	The set of all possible points within a set of constraints
Minimax	An algorithm that finds optimal moves for a player in a two player game by traversing the game tree
Alpha-beta pruning	A heuristic improvement to minimax that doesn't explore fewer promising nodes, hence increasing speed

Monte Carlo Heuristic Tree Search	(MCTS) an advanced game tree search algorithm that employs a weighted random traversal to increase speed in finding optimal moves
Genetic algorithms	Algorithms designed to optimise a model, based on the way that biological systems do, namely genetics and evolution.
Game theory	The study of mathematical models of strategic interaction between rational decision-makers
Convolutional neural networks	(CNNs) neural networks with specialised convolutional layers to improve image processing
Deep neural networks	(DNNs) neural networks with many hidden layers, to allow modelling of complex data sets
Recurrent neural networks	(RNNs) neural networks which link back into themselves to allow simulated memory, and real time processing

9 References

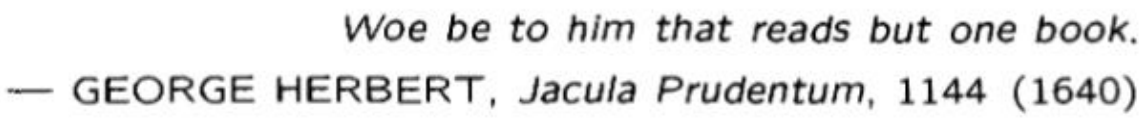


Figure 27: Donald Knuth: The Art of Programming Volume 1: Fundamental Algorithms, page xiv [35]

9.1 Citations

[1] ChessBazaar, [Online]. Available: http://www.chessbazaar.com/blog/wp-content/uploads/2016/10/IMG_9980.jpg. [Accessed 12th May 2019].

[2] Deep Dream Generator, “Deep Dream Generator,” [Online]. Available: https://deepdreamgenerator.com/. [Accessed 12th May 2019].

[3] Merriam-Webster, “Game | Definition of Game by Merriam-Webster,” [Online]. Available: https://www.merriam-webster.com/dictionary/game. [Accessed 30th April 2019].

[4] Astral Castle, “A History of Board Games,” 5th December 2003. [Online]. Available: https://web.archive.org/web/20031205225710/http://www.astralcastle.com/games/index.htm. [Accessed 30th April 2019].

[5] TechCrunch, “Video game revenue tops $43 billion in 2018, and 18% jump from 2017 | TechCrunch,” [Online]. Available: https://techcrunch.com/2019/01/22/video-game-revenue-tops-43-billion-in-2018-an-18-jump-from-2017/. [Accessed 30th April 2019].

[6] K. Jameson, “Game Theory and its Applications,” in *2014 Sr. Seraphim Gibbons Undergraduate Symposium*, Minneapolis, 2013.

[7] A. Rachnog, “Neural networks for algorithmic trading. Correct time series forecasting + backtesting,” A Medium Corporation, 11th May 2017. [Online]. Available: https://medium.com/@alexrachnog/neural-networks-for-algorithmic-trading-1-2-correct-time-series-forecasting-backtesting-9776bfd9e589. [Accessed 30th April 2019].

[8] Y. LeCun, C. Cortes and C. Burges, “MNIST handwritten digit database, Yann LeCun, Corinna Cortes and Chris Burges,” [Online]. Available: http://yann.lecun.com/exdb/mnist/. [Accessed 10th May 2019].

[9] J. Brownlee, “Machine Learning is Popular Right Now,” machinelearningmastery, [Online]. Available: https://machinelearningmastery.com/machine-learning-is-popular/. [Accessed 30th April 2019].

[10] Daffodil Software, “9 Applications of Machine Learning from Day-to-Day Life,” medium.com, 31st July 2017. [Online]. Available: https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0. [Accessed 2nd May 2019].

[11] Nature, “Machine learning - Latest research and news | Nature,” [Online]. Available: https://www.nature.com/subjects/machine-learning. [Accessed 2nd May 2019].

[12] Nuffield Council on Bioethics, “Artificial intelligence (AI) in healthcare and research,” May 2018. [Online]. Available: http://nuffieldbioethics.org/wp-content/uploads/Artificial-Intelligence-AI-in-healthcare-and-research.pdf. [Accessed 2nd May 2019].

[13] R. Munroe, “xkcd: Machine Learning,” 17th May 2017. [Online]. Available: https://imgs.xkcd.com/comics/machine_learning.png. [Accessed 10th May 2019].

[14] M. Cilimkovic, “Neural Networks and Back Propagation Algorithm,” dataminingmasters, [Online]. Available: http://dataminingmasters.com/uploads/studentProjects/NeuralNetworks.pdf. [Accessed 1st May 2019].

[15] ThoughtCo, “neuron-anatomy,” 10th September 2018. [Online]. Available: https://www.thoughtco.com/thmb/wBA889I01ThIHWIjCHAeimx7Q7Y=/1500x0/filters:no_upscale():max_bytes(150000):strip_icc()/neuron-anatomy-58530ffe3df78ce2c34a7350.jpg. [Accessed 2nd May 2019].

[16] R. Bailey, “Human Biology: Neurons and Nerve Impulses,” ThoughtCo., 10th September 2019. [Online]. Available: https://www.thoughtco.com/neurons-373486. [Accessed 2nd May 2019].

[17] F. Rosenblatt, “The perceptron a perveiving and recognizing automaton,” Cornell Aeronautical Laboratory, 1957.

[18] F. Rosenblatt, “The Perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review,* vol. 65, no. 6, pp. 386-408, 1958.

[19] D. J. C. MacKay, “Activity rule (a) ii. Sigmoid (logistic function),” in *Information Theory, Inference, and Learning Algorithms*, Cambridge, Cambridge University Press, 2003, p. 471.

[20] M. A. Nielsen, 2015. [Online]. Available: http://neuralnetworksanddeeplearning.com/images/tikz41.png. [Accessed 2nd May 2019].

[21] R. Rojas, “Neural Networks - A Systematic Introduction,” in *Neural Networks - A Systematic Introduction*, Springer-Verlag, Berlin, New-York, 1996, p. 476.

[22] M. Minsky, Neural Nets and the Brain: Model Problem, Princeton: Unpublished Dissertation, Princeton University, 1954.

[23] Y. LeCun, “A Theoretical Framework for Back-Propagation,” University of Toronto, Toronto.

[24] D. E. Rumelhart, G. E. Hinton and R. J. Williams, “Learning representations by back-propagating errors,” *Nature,* vol. 323, pp. 533-536, 1986.

[25] M. A. Nielsen, “Chapter 2: How the backpropagation algorithm works,” in *Neural Networks and Deep Learning*, Determination Press, 2015.

[26] Tensorflow, “A Neural Network Playground,” tensorflow.org, [Online]. Available: https://playground.tensorflow.org/. [Accessed 10th May 2019].

[27] M. Cook, “It Takes Two Neurons To Ride a Bicycle,” http://citeseerx.ist.psu.edu, Pasadena, CA, 2004.

[28] C. L. Bouton, “Nim, A Game with a Complete Mathematical Theory.,” *Annals of Mathematics; Second Series,* vol. 3, no. 1/4, pp. 35-39, 1901-1902.

[29] Nrich, “Nim-like Games : nrich.maths.org,” nrich.maths.org, [Online]. Available: https://nrich.maths.org/1209. [Accessed 2nd May 2019].

[30] “combinatorial game theory - Neutral tic tac toe - MathOverflow,” MathOverflow, [Online]. Available: https://mathoverflow.net/questions/24693/neutral-tic-tac-toe. [Accessed 1st May 2019].

[31] T. E. Plambeck and G. Whitehead, “The Secrets of Notakto: Winning at X-only Tic-Tac-Toe,” *arXiv,* 8th January 2013.

[32] M. Gardner, “How to build a game-learning machine and then teach it to play and to win,” *The Scientific American,* March 1962.

[33] M. Gardner, “A Matchbox Game-Learning Machine,” gwern.net, [Online]. Available: https://www.gwern.net/docs/rl/1991-gardner-ch8amatchboxgamelearningmachine.pdf. [Accessed 10th May 2019].

[34] W. contributors, “Tic-tac-toe-game-tree - Game Tree - Wikipedia,” 1st April 2007. [Online]. Available: https://en.wikipedia.org/wiki/Game_tree#/media/File:Tic-tac-toe-game-tree.svg. [Accessed 1st May 2019].

[35] D. Knuth, “Trees,” in *The art of computer programming : fundamental algorithms*, Addison-Wesley, 1997, p. 308.

[36] W. contributors, “Minimax,” Wikipedia, The Free Encyclopedia., 1st May 2019. [Online]. Available: https://en.wikipedia.org/wiki/Minimax#Pseudocode. [Accessed 1st May 2019].

[37] “AlphaGo Zero: Learning from scratch,” DeepMind, 19th October 2017. [Online]. Available: https://deepmind.com/blog/AlphaGo-zero-learning-scratch/. [Accessed 30th April 2019].

[38] D. Silver, J. Schrittwieser, I. Antonoglou, K. Simonyan, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre and G. van den Driessche, “Mastering the game of Go without human knowledge,” *Nature,* vol. 550, pp. 354-359, 2017.

[39] C. Metz, “In Two Moves, AlphaGo and Lee Sedol Redefined the Future | Wired,” Wired, 16th March 2016. [Online]. Available: https://www.wired.com/2016/03/two-moves-AlphaGo-lee-sedol-redefined-future/. [Accessed 30th April 2019].

[40] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan and D. Hassabis, “Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm,” *arXiv,* 5th December 2017.

[41] D. R. Hofstadter, “Chunking and Chess Skill,” in *Gödel, Escher, Bach: an Eternal Golden Braid*, The Harvester Press, 1979, p. 286.

[42] S. Gibbs, “AlphaZero AI beats champion chess program after teaching itself in four hours | The Guardian,” The Guardian, 7th December 2017. [Online]. Available: https://www.theguardian.com/technology/2017/dec/07/alphazero-google-deepmind-ai-beats-champion-program-teaching-itself-to-play-four-hours. [Accessed 1st May 2019].

[43] BBC News, “Google's 'superhuman' DeepMind AI claims chess crown - BBC News,” 6th December 2017. [Online]. Available: https://www.bbc.co.uk/news/technology-42251535. [Accessed 1st May 2019].

[44] J. Osborne, “Google's Tensor Processing Unit explained: this is what the future of computing looks like | TechRadar,” TechRadar, 22nd August 2016. [Online]. Available: https://www.techradar.com/news/computing-components/processors/google-s-tensor-processing-unit-explained-this-is-what-the-future-of-computing-looks-like-1326915. [Accessed 1st May 2019].

[45] PeterDoggers, “AlphaZero Chess: Reactions From Top GMs, Stockfish Author,” chess.com, 5th October 2018. [Online]. Available: https://www.chess.com/news/view/alphazero-reactions-from-top-gms-stockfish-author. [Accessed 1st May 2019].

[46] Chessdom, “Komodo MCTS (Monte Carlo Tree Search) is the new star of TCEC | Chessdom,” chessdom.com, 18th December 2018. [Online]. Available: http://www.chessdom.com/komodo-mcts-monte-carlo-tree-search-is-the-new-star-of-tcec/. [Accessed 1st May 2019].

[47] J. Schaeffer, “Chinook,” 2nd August 2004. [Online]. Available: https://web.archive.org/web/20040930164251/http://www.cs.ualberta.ca/~chinook/. [Accessed 6th May 2019].

[48] J. Schaeffer, “Marion Tinsley: Human Perfection at Checkers?,” wylliedraughts.com, [Online]. Available: http://www.wylliedraughts.com/Tinsley.htm. [Accessed 6th May 2019].

[49] J. Schaeffer, N. Burch, Y. Björnsson, A. Kishimoto, M. Müller, R. Lake, P. Lu and S. Sutphen, “Checkers Is Solved,” *Science,* vol. 317, pp. 1518-1522, 2007.

[50] G. Van Rossum, B. Warsaw and N. Coghlan, “PEP 8 -- Style Guide for Python Code,” The Python Software Foundation, 5th July 2001. [Online]. Available: https://www.python.org/dev/peps/pep-0008/. [Accessed 30th April 2019].

[51] D. Dua and C. Graff, “UCI Machine Learning Repository,” 2019.

[52] W. contributors, “Alpha–beta pruning,” Wikipedia, The Free Encyclopedia., 26th April 2019. [Online]. Available: https://en.wikipedia.org/wiki/Alpha%E2%80%93beta_pruning#Pseudocode. [Accessed 10th May 2019].

[53] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks fromOverfitting,” *Journal of Machine Learning Research,* vol. 15, pp. 1929-1958, 2014.

[54] E. J. Crowley, J. Turner, A. Storkey and M. O'Boyle, “Pruning neural networks: is it time to nip it in the bud?,” *arXiv,* 19th January 2019.

[55] R. Vasudev, “How to Initialize weights in a neural net so it performs well?,” Hackernoon, 29th May 2018. [Online]. Available: https://hackernoon.com/how-to-initialize-weights-in-a-neural-net-so-it-performs-well-3e9302d4490f. [Accessed 10th May 2019].

[56] M. Thoma, “Analysis and Optimization of Convolutional Neural NetworkArchitectures,” 2017.

[57] S. V. Avinash, “Understanding Activation Functions in Neural Networks,” medium.com, 30th March 2017. [Online]. Available: https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0. [Accessed 6th May 2019].

[58] Y. LeCun, O. Matan, B. Boser, J. S. Denker, D. Henderson, H. R. E, W. Hubbard, L. D. Jackel and H. S. Baird, “Handwritten Zip Code Recognition with Multilayer Networks,” AT&T Bell Laboratories, Holmdel, New Jersey, 1990.

[59] A. Krizhevsky, I. Sutskever and G. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *NIPS 2012: Neural Information Processing Systems*, Lake Tahoe, Nevada, 2012.

[60] I. Sutskever, “Training Recurrent Neural Networks,” Toronto, 2013.

[61] A. Krizhevsky, I. Sutskever and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, 2012.

[62] W. contributors, “Logistic-Curve - Logistic Function - Wikipedia,” 2nd July 2008. [Online]. Available: https://en.wikipedia.org/wiki/Logistic_function#/media/File:Logistic-curve.svg. [Accessed 1st May 2019].

[63] E. R. Berlekamp, J. H. Conway and R. K. Guy, “Green Hackenbush, The Game of Nim, and Nimbers & Getting Nimble with Nimbers,” in *Winning Ways for Your Mathematical Plays, Volume 1*, Wellesley, MA, AK Petes Ltd, 2001, pp. 40-43.

[64] J. Han and C. Moraga, “The influence of the sigmoid function parameters on the speed of backpropagation learning,” *In: Mira J., Sandoval F. (eds) From Natural to Artificial Neural Computation. IWANN 1995. Lecture Notes in Computer Science, vol 930,* vol. 930, 1995.