The rubric for this examination should read as follows

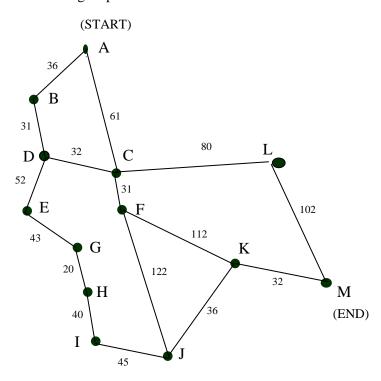
Marks will be awarded for the first FOUR questions in the answer book. Clearly cross through any questions that you do NOT wish to be considered and ensure you state on the front of the answer book the FOUR questions you have attempted.

Question 1

a) Using an example, show how Uniform Cost Search operates?

(8 marks)

b) Consider the following map.



Using the A* algorithm work out a route from town A to town M. Use the following cost functions.

- \triangleright G(n) = The cost of each move as the distance between each town (shown on map).
- \rightarrow H(n) = The Straight Line Distance between any town and town M. These distances are given in the table below.

Provide the search tree for your solution and indicate the order in which you expanded the nodes. Finally, state the route you would take and the cost of that route.

Straight Line Distance to M

A	223
В	222
С	166
D	192

Е	165
F	136
G	122
Н	111

Ι	100
J	60
K	32
L	102

M	0

In your answer provide the following

i) The search tree that is produced, showing the cost function at each node

(10 marks)

ii) State the order in which the nodes were expanded

(2 marks)

State the route that is taken, and give the total cost

(1 mark)

Explain the relationship between the A* algorithm and the Uniform Cost Search algorithm?

c)

(4 marks)

Question 2

a) For each of the truth tables below say whether it is possible for a perceptron to learn the required output.

In each case, explain the reason behind your decision.

i)	Input	0	0	1	1
	Input	0	1	0	1
	Required Output	1	0	0	1

ii)	Input	0	0	1	1
	Input	0	1	0	1
	Required Output	1	1	0	0

iii)	Input	0	0	1	1
	Input	0	1	0	1
	Required Output	1	1	1	1

(9 marks)

b) A perceptron with two inputs has a threshold level set at the point at which it will fire (i.e. output a one). It is sometimes convenient to always set the threshold level to zero.

Show how this can be achieved by describing two perceptrons which act in the same way but one has its threshold set to a non-zero figure and the other perceptron has a zero threshold.

(13 marks)

c) Why might it be a good idea to build a perceptron with a zero threshold figure?

(3 marks)

Question 3

Write brief notes on each of the following

a) The Chinese Room. Marks will be awarded for a description of The Chinese Room, what it refutes and some of the criticisms/replies for The Chinese Room.

(9 marks)

b) Game Playing and Checkers. Marks will be awarded or describing the history of checkers and citing relevant literature.

(8 marks)

c) McCulloh and Pitts Neural Network. Marks will be awarded for describing a McCulloch and Pitts network, the main attributes and for giving an example.

(8 marks)

Question 4

a) With reference to the Travelling Salesman Problem explain what is meant by combinatorial explosion and what effect it has in finding an optimal solution?

(5 marks)

- b) Match the following AI pioneers to the area of AI for which they are associated with.
 - 1. McCarthy
 - 2. Shannon
 - 3. Minsky and Papert
 - 4. Newell and Simon
 - 5. Samuel
 - 6. Schaeffer
 - 7. Searle
 - 8. McCulloch and Pitts
 - 9. Shortliffe
 - 10. Weizenbaum

- a) Seminal work for checkers
- b) First Neural Network
- c) Chinook
- d) Proved that perceptrons cannot learn functions which are not linearly separable
- e) Chinese Room
- f) ELIZA
- g) Physical Symbol System Hypothesis
- h) Seminal work for chess
- i) LISP
- j) MYCIN

(10 marks)

c) What advances do you think need to be made in order for The Turing Test to be passed? (No marks will be awarded for describing The Turing Test.)

(10 marks)

Question 5

For the 8-puzzle problem, given this initial starting state

2	8	7
3	1	4
5	6	

a) Using breadth first search, show the search tree that would be built down to level 2 (assume level zero is the root of the tree).

(5 marks)

b) Using depth first search, show the state of the search tree down the level 3 (stop once you have expanded one node that goes to level 3)

(5 marks)

c) What is the worst-case time and space complexity of the above two algorithms.

(5 marks)

d) Describe the terms complete and optimal with regards to evaluating search strategies?

(4 marks)

e) Are either depth-first-search or breadth-first-search complete or optimal? Justify your answer.

(6 marks)

Question 6

(a) Nim is a two-player game The rules are as follows.

The game starts with a single stack of 7 tokens. At each move a player selects one stack and divides it into two non-empty, non-equal stacks. A player who is unable to move loses the game.

Draw the complete search tree for nim.

(10)

(b) Assume two players, min and max, play nim (as described above). Min plays first.

If a terminal state in the search tree developed above is a win for min, a utility function of zero is assigned to that state. A utility function of 1 is assigned to a state if max wins the game.

Apply the minimax algorithm to the search tree to assign utility functions to all states in the search tree.

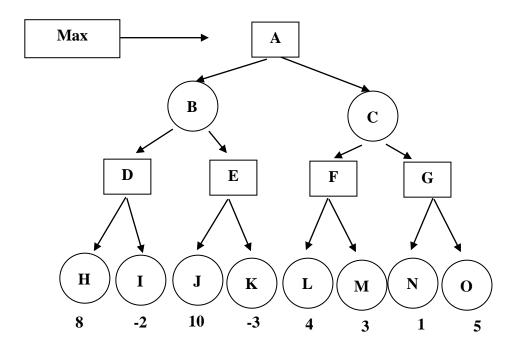
(3)

(c) If both min and max play a perfect game, who will win? Explain your answer and show the path taken by the player that wins.

(3)

(d) Given the following search tree, apply the alpha-beta pruning algorithm to it and show the search tree that would be built by this algorithm. Make sure that you show where the alpha and beta cuts are applied and which parts of the search tree are pruned as a result.

(9)



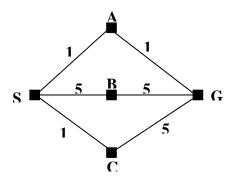
Question 1 – Model Answer

a) Using an example, show Uniform Cost Search operates?

Breadth first search finds the shallowest goal state and that this will be the cheapest solution so long as the path cost is a function of the depth of the solution. But, if this is not the case, then breadth first search is not guaranteed to find the best (i.e. cheapest) solution. Uniform cost search remedies this.

It works by always expanding the lowest cost node on the fringe, where the cost is the path cost, g(n). In fact, breadth first search is a uniform cost search with g(n) = DEPTH(n).

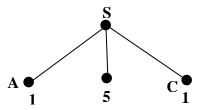
Consider the following problem.



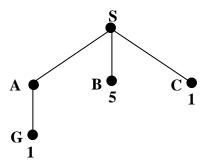
We wish to find the shortest route from S to G; that is S is the initial state and G is the goal state. We can see that SBG is the shortest route but if we let breadth first search loose on the problem it will find the path SAG; assuming that A is the first node to be expanded at level 1.

But this is how uniform cost search tackles the problem.

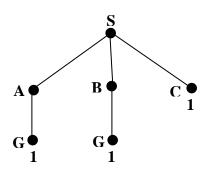
We start with the initial state and expand it. This leads to the following tree



The path cost of A is the cheapest, so it is expanded next; giving the following tree



We now have a goal state but the algorithm does not recognize it yet as there is still a node with a cheaper path cost. In fact, what the algorithm does is order the queue by the path cost so the node with cost 11 will be behind node B in the queue. Node B (being the cheapest) is now expanded, giving



A goal state (G) will now be at the front of the queue. Therefore the search will end and the path SBG will be returned.

In summary, uniform cost search will find the cheapest solution provided that *the cost of the path never decreases as we proceed along the path*. If this requirement is not met then we never know when we will meet a negative cost. The result of this would be a need to carry out an exhaustive search of the entire tree.

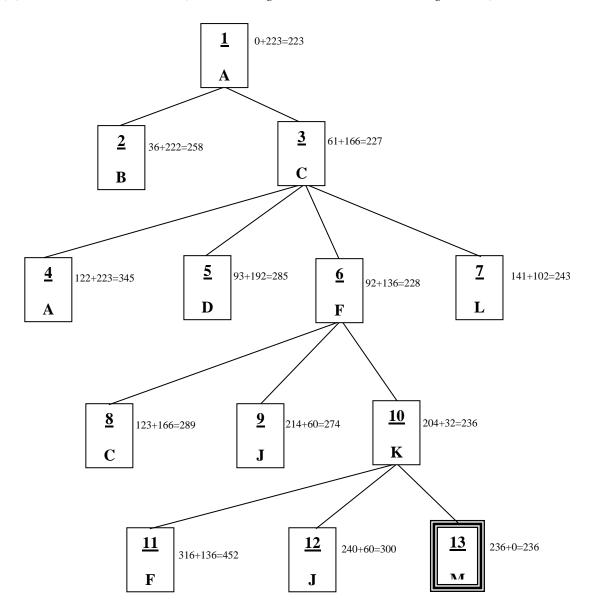
b) Using A* Algorithm

The Search Tree

The figures next to each node represent the G(n) and H(n) functions, where

G(n) = The cost of the search so far (i.e. distance travelled)

H(n) = The heuristic value (i.e. the straight line distance to the target town)



The route you would take is A, C, F, K, M at a cost of 236.

c) Explain the relationship between the A* algorithm and the Uniform Cost Search algorithm?

Uniform cost search minimises the cost of the path so far, g(n). Uniform cost search is both optimal and complete but can be very inefficient.

We combine this evaluation function, g(n), with a heuristic evaluation, h(n) to give the evaluation function for the A^* algorithm, i.e.

$$f(n) = g(n) + h(n)$$

As g(n) gives the path cost from the start node to node n and h(n) gives the estimated cost of the cheapest path from n to the goal, we have

f(n) = estimated cost of the cheapest solution through n

A* is both optimal and complete is the heuristic is admissable (i.e. does not overestimate the cost to the goal). It also, in the worst case, has the same time and space complexity as uniform cost search.

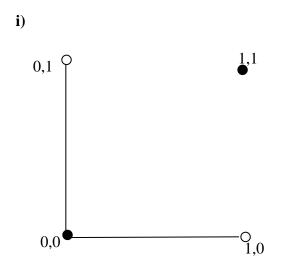
Question 2 – Model Answer

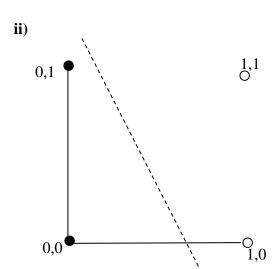
a) For each of the truth tables [not shown in the model answer] below say whether it is possible for a perceptron to learn the required output.

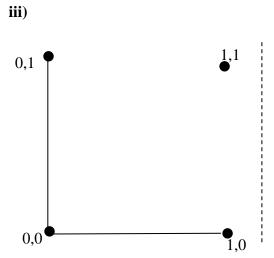
Three marks for each

Only i) cannot be learnt. This is because it is not *linearly separable*. This can be shown on the diagrams below (where the outputs have been plotted and the filled circles represent a 1 and the hollow circles represent a zero.).

Those problems which are linearly separable can have a line dividing the "1" outputs from the "0" outputs. In the case of i) this is not possible.







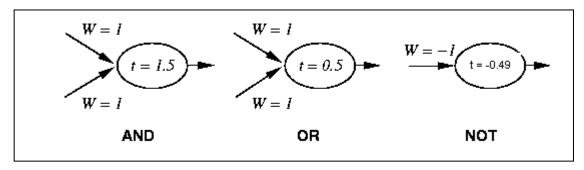
b) A perceptron with two inputs has a threshold level set at the point at which it will fire (i.e. output a one).

It is sometimes convenient to always set the threshold level to zero.

Show how this can be achieved by describing two perceptrons which act in the same way but one has its threshold set to a non-zero figure and the other perceptron has a zero threshold.

Perceptrons with two inputs and the threshold non-zero

A perceptron (with two inputs) to act as a logic gate could be modelled as follows (three examples are shown – students would only need to show one example).



If we consider the AND function we can see that it acts correctly for the four possible inputs (see table below).

Note in the table Sum is defined as (Input 1 * Weight 1) + (Input 2 * Weight 2)

Step(t) is defined as returning 1 if Sum => t else 0 and the output of Step(t) is also the output (or activation level) of the perceptron.

<u>AND</u>

Input 1	Input 2	Weight 1	Weight2	Sum	Step(t)
0	0	1	1	0	0
0	1	1	1	1	0
1	0	1	1	1	0
1	1	1	1	2	1

Similarly, OR and NOT can be shown as follows

<u>OR</u>

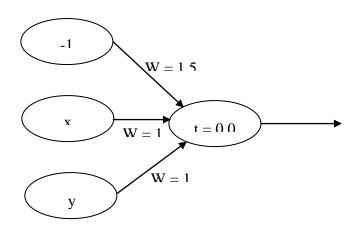
Input 1	Input 2	Weight 1	Weight 2	Sum	Step(t)
0	0	1	1	0	0
0	1	1	1	0.5	1
1	0	1	1	0.5	1
1	1	1	1	1	1

NOT

Input 1	Weight 1	Sum	Step(t)
0	-1	0	1
1	01	-0.49	0

Perceptron with three inputs and the threshold set to zero

It is possible to have an extra input whose activation is set to -1 and the weight from that input unit to the output neuron is set to the required threshold level. Diagramatically this can be shown as follows



It can be shown that this acts in the same way as the previous perceptrons.

To demonstrate the perceptrons act in the same way the following tables are given

AND (weight on extra neuron = 1.5)

Input 1	Input 2	Input 3	Sum	Step(0)
-1	0	0	-1.5	0
-1	0	1	-0.5	0
-1	1	0	-0.5	0
-1	1	1	0.5	1

OR (weight on extra neuron = 0.5)

Input 1	Input 2	Input 3	Sum	Step(0)
-1	0	0	-0.5	0
-1	0	1	0	1
-1	1	0	0	1
-1	1	1	0.5	1

NOT (weight on extra neuron = -0.49)

Input 1	Input 2	Sum	Step(0)
-1	0	0.49	1
-1	1	-0.51	0

Again, I would not expect the students to show three examples. They are just shown for completeness.

c) Why might it be a good idea to build a perceptron with a zero threshold figure?

When learning, the algorithm only has to adjust weights and not thresholds and weights.

Question 3 – Model Answer

This questions is designed to give students who have done general revisions, some chance of picking up marks.

a) The Chinese Room

(9 marks)

Marks will be awarded for covering the following information.

Description

The system comprises a human, who only understands English, a rule book, written in English, and two stacks of paper. One stack of paper is blank. The other has indecipherable symbols on them.

In computing terms the human is the CPU, the rule book is the program and the two stacks of paper are storage devices.

The system is housed in a room that is totally sealed with the exception of a small opening.

The human sits inside the room waiting for pieces of paper to be pushed through the opening. The pieces of paper have indecipherable symbols written upon them.

The human has the task of matching the symbols from the "outside" with the rule book. Once the symbol has been found the instructions in the rule book are followed. This may involve writing new symbols on blank pieces of paper or looking up symbols in the stack of supplied symbols.

Eventually, the human will write some symbols onto one of the blank pieces of paper and pass these out through the opening.

Assume that the symbols passed into the room were valid Chinese sentences, which posed questions. Also assume that pieces of paper passed out of the room were also valid Chinese sentences, which answered those questions. Searle argues that we have a system that is capable of passing the Turing Test and is therefore intelligent according to Turing. But Searle also argues that the system does not understand Chinese as it just comprises a rule book and stacks of paper which do no understand Chinese.

4 marks (prorata)

What it Demonstrates

It refutes the Turing Test, saying that The Chinese Room passes the Turing Test but it does not understand. Therefore, running the right program does not necessarily generate understanding.

2 marks

Criticisms of The Chinese Room

There were various arguments against Searle's conclusion. One of them Searle called the Systems Reply. This argued that the system as a whole understand Chinese.

Searle's argument was to remove all the elements the system and place the entire system inside the brain of the human. Therefore, the human would memorise all the rules and the stacks of paper.

Searle's now argues that the only thing that can understand Chinese is the human and if you asked the human (in English) if they understood Chinese the answer would be no.

There were other arguments

- The Robot Reply argues we could internalise everything inside a robot (android) so that it appears like a human. Searle argues that nothing has been achieved by adding motors and perceptual capabilities.
- The Brain Simulator Reply argues we could write a program that simulates the brain (neurons firing etc.) Searle argues we could emulate the brain using a series of water pipes and valves. Can we now argue that the water pipes understand? He claims not.
- The Combination Reply argues we could combine all three of the previous suggestions..

 Searle argues that this does not progress the argument. As we know how the android operates Searle claims it is not carrying out strong AI

3 marks (prorata)

b) Game Playing and Checkers

Learning techniques were being used for checkers (draughts) as far back as the 1950's with Samuel's seminal work. Arthur Samuel, in 1952 wrote the first checkers program. The original program was written for an IBM 701 computer. In 1954 he re-wrote the program for an IBM 704 and added a learning mechanism. What makes this program stand out in AI history is that the program was able to learn its own evaluation function. Taking into account the IBM 704 had only 10,000 words of main memory, magnetic tape for long-term storage and a cycle time of almost one-millisecond, this can be seen as a major achievement in the development of AI. Samuel made the program play against itself and after only a few days play, the program was able to beat its creator and compete on equal terms with strong human opponents.

It remains as a testament to Samuel that there was little more work done on checkers until Jonathon Schaeffer et. al. developed Chinook.

Jonathan Schaeffer's work led to the development of Chinook. In 1992 Chinook won the US Open and subsequently challenged for the world championship. Dr. Marion Tinsley had been the world champion for over 40 years. In that time he had only lost three games. Playing Chinook, he lost his fourth and fifth game but ultimately won the match by 21.5 points to Chinook's 18.5 points. In August 1994 there was a re-match but the match ended prematurely when Dr. Tinsley had to withdraw for health reasons. As a result of this Chinook become the official world champion. claimed that Chinook was rated at 2814. The best human players are rated at 2632 and 2625.

Despite Chinook becoming the world champion, the search has continued for a checkers player that is built using "true" AI techniques. Chellapilla and Fogel have developed Anaconda, named, due to the strangle hold it places on its opponent. It is also called Blondie24 [FOG01] which is the name it used when playing on the internet This name was chosen in a successful attempt to try and attract players on the assumption they were playing against a blonde 24 year old female. Anaconda (or Blondie24) uses an artificial neural network (ANN), with 5000 weights, which are evolved by an evolutionary strategy. The inputs to the ANN are the current board position and it outputs a value which is used in a mini-max search. During the training period the program is given no information other than whether it won or lost (it is not even told by how much). Anaconda is certainly not given any strategy and contains no database of opening and ending game positions. Co-evolution is used to develop Anaconda, by playing games against itself. Once Anaconda is able to play at a suitable level, it often searches to a depth of 10, but depths of 6 and 8 are also common in play.

I am not necessarily looking for this much detail. Just evidence they know some of the history and can cite some of the literature.

8 marks (prorata, with 2 of them for citing some literature)

c) McCulloh and Pitts Neural Network

McCulloch & Pitts (McCulloch, 1943) are generally recognised as being the designers of the first neural network. They recognised that combining many simple processing units together could lead to an overall increase in computational power.

Many of the ideas they suggested are still in use today. For example, the idea that a neuron has a threshold level and once that level is reached the neuron fires is still the fundamental way in which artificial neural networks operate.

2 marks (prorata)

We can make the following statements about a McCulloch-Pitts network

The activation of a neuron is binary. That is, the neuron either fires (activation of one) or does not fire (activation of zero).

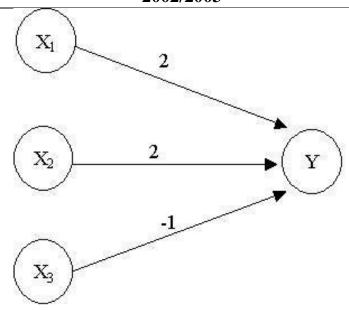
For the network shown below the activation function for unit Y is

$$f(y in) = 1$$
, if $y in >= t$ else 0

where y_in is the total input signal received t is the threshold for Y.

- Neurons is a McCulloch-Pitts network are connected by directed, weighted paths.
- If the weight on a path is positive the path is excitatory, otherwise it is inhibitory.
- All excitatory connections into a particular neuron have the same weight, although different weighted connections can be input to different neurons.
- Each neuron has a fixed threshold. If the net input into the neuron is greater than the threshold, the neuron fires.
- The threshold is set such that any non-zero inhibitory input will prevent the neuron from firing.
- It takes one time step for a signal to pass over one connection.

4 marks for describing main attributes (prorata)



A sample McCulloch-Pitts network is shown above and some of the statements can be observed. In particular, note that the threshold for Y has equal 4 as this is the only value that allows it to fire, taking into account that a neuron cannot fire if it receives a nonzero inhibitory input.

2 marks for giving an example (a logic function would do (AND/OR etc.)) (prorata)

Question 4 – Model Answer

a) With reference to the Travelling Salesman Problem explain what is meant by combinatorial explosion and what effect this has in finding an optimal solution?

The number of solutions is n! (n factorial), where n is the number of cities. This results in an exponential rise in the number of solutions. For example, for 10 cities the number of possible routes is 3,628,800.

This is known as combinatorial explosion where the number of solutions rises exponentially.

The effect this has with regards to TSP is that is quickly becomes impossible to search the entire search space (i.e. enumerate all possible solutions and choose the best route).

Therefore, heuristic methods are often used to find solutions to these problems.

5 marks: prorata

- b) Match the following AI pioneers to the area of AI for which they are associated with.
 - 1. Samuel
 - 2. McCulloch and Pitts
 - 3. Schaeffer
 - 4. Minsky and Papert
 - 5. Searle
 - 6. Weizenbaum
 - 7. Newell and Simon
 - 8. Shannon
 - 9. McCarthy
 - 10. Shortliffe

- a) Seminal work for checkers
- b) First Neural Network
- c) Chinook
- d) Proved that perceptrons cannot learn functions which are not linearly separable
- e) Chinese Room
- f) ELIZA
- g) Physical Symbol System Hypothesis
- h) Seminal work for chess
- i) LISP
- j) MYCIN

OR

1i 2h 3d 4g 5a 6c 7e 8b 9j 10f

10 marks in total (1 mark for each correct answer) 10 marks may seem a lot for this question but if they get one wrong, that means they have at least two wrong AND, part c of this question is not that easy to get full marks.

c) What advances do you think need to be made in order for The Turing Test to be passed?

I am looking for a clear description of the main areas that would need to be advanced.

The three main areas would have to be

- Natural language processing so that the computer can interpret and respond to any question.
- The knowledge base that would have to be held
- How that knowledge would be represented
- How the knowledge would be accessed

The student may also give examples of other areas that would have to be in place.

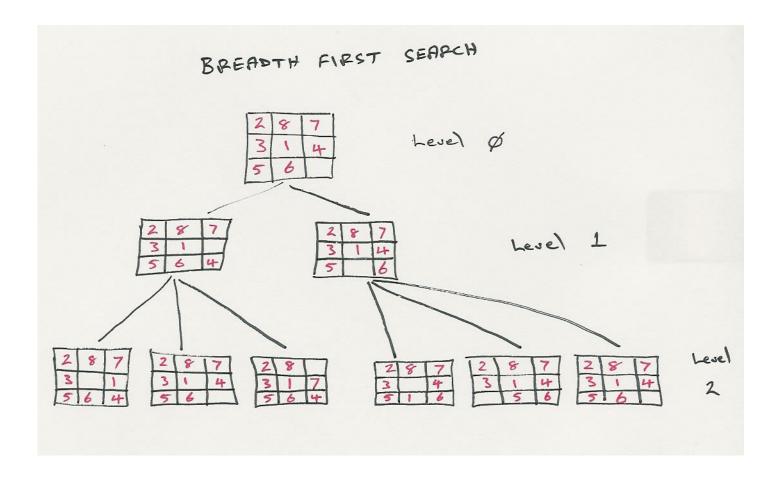
I am not looking for a description of The Turing Test. The student might give a brief one but no marks will be awarded for this.

7 marks – prorata, with 3 available for coming up with their own views/ideas

Question 5 – Model Answer

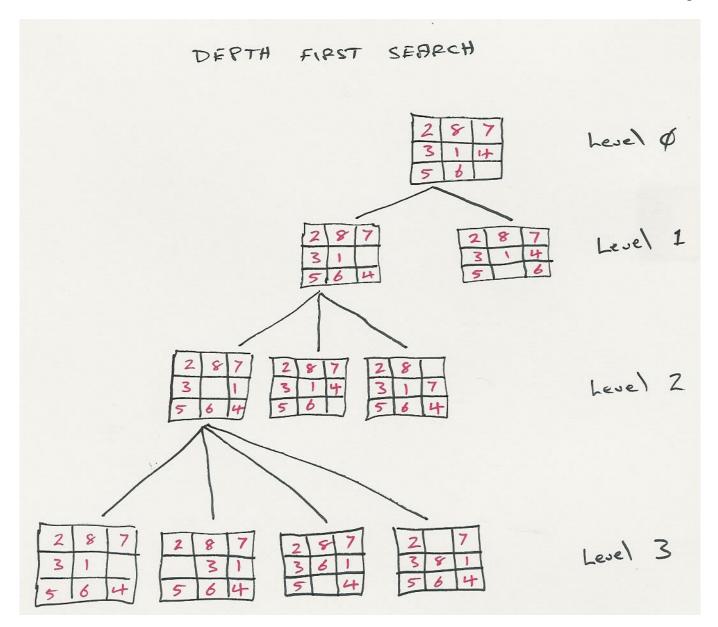
a) Using breadth first search, show the search tree that would be built down to level 2 (assume level zero is the root of the tree).

(5 marks : prorata)



b) Using depth first search, show the state of the search tree down the level 3 (stop once you have expanded one node that goes to level 3)

(5 marks: prorata)



c) What is the worst-case time and space complexity of the above two algorithms.

(1 mark for each part that is correct and 1 mark or giving the definition of the variables)

Evaluation	Breadth First	Depth First
Time	B^{D}	B^{M}
Space	B^{D}	BM

Where B = Branching Factor

D = Depth of Solution

M = Maximum Depth of the Search Tree

d) Describe the terms complete and optimal with regards to evaluating search strategies?

(2 marks each)

Complete: Is the search guaranteed to find a solution if there is one?

Optimal : Is the search guaranteed to find the optimal (cheapest) solution (however cheapest has been

defined for that search problem)?

e) Are either depth-first-search or breadth-first-search complete or optimal? Justify your answer.

Breadth-First Search is *both* **optimal** and **complete** (1 mark for each, i.e. 2 marks available) Depth-First Search is *neither* **optimal** nor **complete** (1 mark for each, i.e. 2 marks available)

1 mark for describing why BFS is optimal and complete. That is, it is systematic and it searches each level and, ultimately it will expand the entire search tree so will find a solution if one exists. Note, BFS is only optimal if the path cost is a function of the search depth. I am not looking for this – so no extra marks.

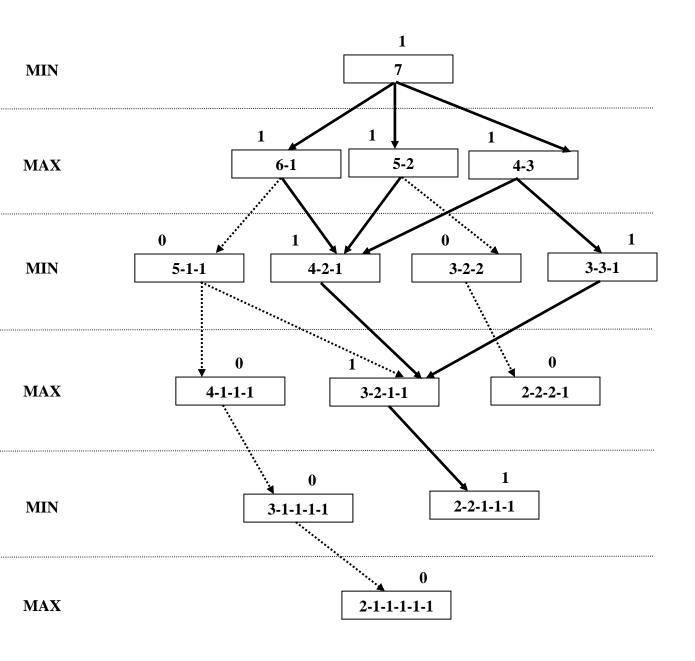
1 mark for describing why DFS is not optimal or complete. That is, it can disappear into an infinite search down one particular branch of the tree and there is no guarantee that there is a solution down that path. No marks for talking about Depth Limited Search or Iterative Deepening Search – as I have not asked for these.

Question 6 – Model Answer

(a) Draw the complete search tree for nim.

The search tree is shown below. The student might draw a search tree which duplicates some of the nodes. This is acceptable.

Note: At this stage, only the search tree is required. The "Min", "Max" and the Utility Functions are not required.



(b) Assume two players, min and max, play nim (as described above). Min plays first.

If a terminal state in the search tree developed above is a win for min, a utility function of zero is assigned to that state. A utility function of 1 is assigned to a state if max wins the game.

Apply the minimax algorithm to the search tree to assign utility functions to all states in the search tree.

The utility functions are shown in the search tree above. The student should initially assign values to the terminal states and then propagate these up the tree, depending on whether it is min or max's turn to move (min minimises the utility function of its child nodes, max maximises).

The student should mark the tree with the players turn – deduct one mark if this is not done. The other two marks are for correctly assigning the correct utility vales to each node. Give one mark if the student appears to be doing it correctly but has made a mistake.

(c) If both min and max play a perfect game, who will win? Explain your answer and show the path taken by the player that wins.

As the value at the top of the search tree is 1, it shows that maximise is guaranteed to win (if it plays a perfect game). This is because minimise will never be able to follow a path that leads to a utility function of 0. One mark for pointing this out.

Two marks for showing the path that would be taken, if both players played a perfect game. These are shown in bold in the search tree above.

(d) Given the following search tree, apply the alpha-beta pruning algorithm to it and show the search tree that would be built by this algorithm. Make sure that you show where the alpha and beta cuts are applied and which parts of the search tree are pruned as a result.

The search tree that the student should re-produce is shown below. An example very similar to this was shown in the course notes (it has the same structure, different values).

If the student produces this search tree they should receive 6 marks (note any extra nodes they produce should be penalised as the idea behind the alpha-beta algorithm is that it restricts the number of nodes that are expanded).

The other three marks are to be allocated based on the way the student explains how the search tree was built. One mark for stating that alpha-beta pruning *must* use Depth First Search.

One mark for explaining why the alpha cut off can be made. One mark why the beta cut off can be made.

This is the explanation given in the course notes

Assume we have evaluated, using depth first search down to node I. This means we can maximise node D to a value of 8. If we now continue the search we will eventually evaluate node J and assign it a value of 10. At this point we do not need to evaluate K (or any of its siblings); for the following reasons.

- Node E already has a value of at least 10 (as MAX is trying to maximise this value so it cannot take on a smaller value).
- Node E is already greater than node D and, as MIN will be trying to minimise these two values, it is always going to choose D
 over E
- Therefore, we can cut off the search at this point (shown an the diagram).

Similarly, we can cut of the search from node G downwards.

- Having evaluated F to 4, C cannot be more than this value (as C is trying to minimise).
- B is already greater than 4.

• Therefore, MAX is always going to prefer B over C, so we cannot cut off the rest of the search below C.

