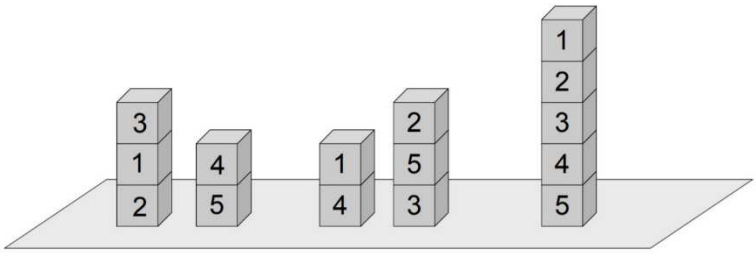
Exam with Consultation. Duration: 1h30m. Normal Exam

Exercise 1) Search Problems (5 Values)

Suppose the well-known blocks world represented in the figure. Imagine that the goal is to reach the solution shown in the right of the figure (where the 5 blocks are stacked with the highest value on the bottom and the lowest on the top). States 1 and 2 represent possible initial states. Imagine that in each operation you can move a block to the floor or on top of a block of higher value than the block moved. Suppose that, at any given moment, you can only have a maximum of 3 stacks of blocks. The ultimate goal is to get all the blocks to be stacked as shown in the right of the figure.



1.1) Formulate the problem as a search problem.

1.2) Calculate the problem state-space size. What is the mean ramification factor of the problem? Justify.

1.3) Define an admissible heuristic that allows to apply the A * algorithm to the problem and present the pseudo-code to calculate it.

1.4) Assuming the figure left state as the initial state, present the trees for solving the problem using the breadth-first search algorithm and the A * algorithm with the Heuristic defined in the previous question.

1.5) Do you think it is possible to apply bidirectional search to the problem? Justify explaining how it could be applied or why it cannot be applied.

Exercise 2) Optimization (5 Values)

A governmental entity with competence in the area of food inspection needs to carry out periodic inspections at different establishments. For this purpose, an inspection brigade is available, for which an inspection route is generated on a daily basis. The execution of the route must take place within a working day (7 hours). When generating the route, it is necessary to take into account the traveling time and the time spent on each inspection, at the location of the establishment in question. Given a set of n establishments, duly georeferenced (for simplification consider a pair of coordinates (x, y) in km), one intends to generate the inspection route for a single day, seeking to maximize the number of inspected establishments. For simplification, suppose that for each establishment there is an inspection time (between half an hour and 3 hours) and that the travel time, in hours, from one point to another is equal to 1/10 of the Manhattan distance between those points (for example, between points (3,3) and (13,13) the travel time will be 2 hours.

2.1) Suggest a representation for the files containing instances of this problem to solve. Suggest also a representation for a solution to this problem, where the ordered establishments to be visited by the brigade must be considered, together with the establishments that will not be inspected.

2.2) Identify what is the strong constraint of the problem and what is the problem optimization criterion. Explain how the solution is interpreted and evaluated, exemplifying, and presenting a pseudo-code for calculating the respective evaluation function.

2.3) In order to apply a local search algorithm (for example, simulated annealing), define a neighborhood function and display the respective pseudo-code.

2.4) Present the pseudo-code to solve the problem using simulated annealing. Suppose an initial temperature of 10 and a temperature decrease of 5% per iteration and a maximum of 10000 iterations. The algorithm must also stop after 500 iterations with no improvement in the best solution.

Exercise 3: Artificial Intelligence (10 values).

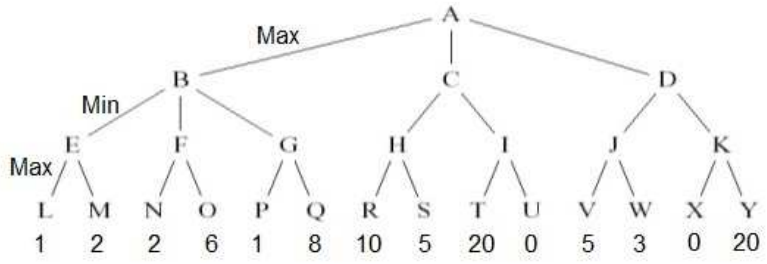
Note: Please solve only 6 of the following 9 questions

9. Present the PEAS description of a home vacuum cleaner robot describing in detail the environment characteristics. Do you think it is possible to use a simple reactive agent to control a robot of this type? Justify.

10. Suppose you apply breadth-first search and iterative deepening algorithms to a problem with a branch expansion factor of the search tree of r = 4 where the solution is at level 5 (exactly at the last node analyzed at level 5). Calculate how many nodes will be generated/analyzed by each of the two methods, calculating the increment of analyzed nodes, in percentage, of the iterative deepening method in relation to the breadth-first search. Comment on the value obtained.

11. Do you think it is possible to create a simple reflex agent (simple reactive agent) to play the game of Go with the guarantee that this agent never loses in this game? Please Justify. What kind of approach do you think is the most appropriate to create a Go playing agent?

12. Apply the minimax algorithm with alpha-beta cuts to the following tree, assuming MAX plays first, then MIN and again MAX, indicating:



a) What is the final value of Nodes A, B, C and D? (Note: since alpha-beta cuts are used, the value may be a number, eg 8 or a condition eg: >=20 or <=8)

b) Which leaf nodes (from L to Y) are not evaluated by the minimax algorithm with alpha-beta cuts?

13. In genetic algorithms, explain how the value of a chromosome is taken into account in the roulette-based selection method.

14. Suppose the following training (individuals 1 to 13) and test sets (individuals 14, 15 and 16). Without the need to do any math, assuming the Euclidean distance, indicate, justifying, which classes are foreseen for individuals 14, 15 and 16 using the Nearest Neighbor algorithm, versions 1NN and 3NN.

Treino	Par1	Par2	Par3	Classe
Ind1	3	8	9	A
Ind2	4	7	8	A
Ind3	2	2	5	B
Ind4	1	3	6	B
Ind5	8	6	5	C
Ind6	9	5	6	C
Ind7	2	1	6	B
Ind8	1	3	5	B
Ind9	4	9	8	A
Ind10	3	8	7	A
Ind11	9	4	6	C
Ind12	8	5	5	C
Ind13	1	1	1	C

Teste	Par1	Par2	Par3	Classe 1NN	Classe 3NN
Ind14	1	1	3	?	?
Ind15	2	4	6	?	?
Ind16	5	9	9	?	?

15. Suppose you have a CSV file with data related to the size (in m2), location (municipality), year of construction (integer value), and type of finishes (scale with 5 possible values) of 10,000 Portuguese properties. For each property, a classification of its "type for sale" is also available in the CSV (basic, medium, luxury, and super-luxury). Suppose you want to use supervised learning to train a classifier to be able to classify the "type for sale" of new properties from their size, location, year, and type of finishes. Present the essential code (summarized), using Python and the Pandas and SciKit Learn libraries, to read the CSV, train a classifier with neural networks, using 5 fold cross-validation, and present the respective confusion matrix and accuracy. Note: You do not need to present the absolutely correct code but only an approximation.

16. In a reinforcement learning problem, upon reaching a state, an agent has four actions at its disposal, whose Q values are as follows: Q (A) = 5, Q (B) = 5, Q (C) = 1, Q (D) = 6. Following an epsilon-greedy action selection policy, what is the probability, rounded to 2 decimal places, of choosing action B for an epsilon value of 0.1?

17. The text in this question has a vocabulary size of 23 (number of different alphanumeric character sequences). Compute the probability of the word "text" in this text, using Laplace smoothing.