Introduction to Artificial Intelligence (IAI)
BSc and MSc Exam

IT University of Copenhagen

29 May 2020

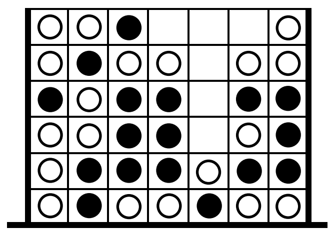
This exam consists of 8 numbered pages with 3 problems. Each problem is marked with the weight in percent it is given in the evaluation. **Notice that Bachelor (BSc) and Master (MSc) students must solve the same problems**.

- Ensure that you have a quiet, comfortable, and undisturbed work environment.
- Use notations and methods that have been discussed in the course.
- Make appropriate assumptions when a problem has been incompletely specified.

Good luck!

Problem 1: Adversarial Search (30%)

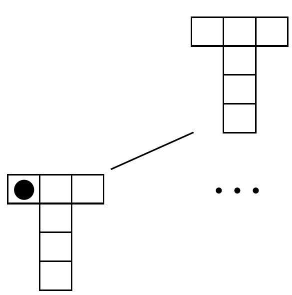
Connect 4 is a two-player board game in which the players first choose a color and then take turns dropping one colored disc from the top into a seven-column, six-row vertically suspended grid. The discs fall straight down, occupying the lowest available space within the column. The objective of the game is to be the first to form a horizontal, vertical, or diagonal line of four of one's own discs. An example of a game state is shown below.



a) Consider building a complete game tree from the initial state of Connect 4, where the grid is empty. How many nodes are there at depth six in this tree?

Assume that MAX plays the black discs and that it is MAX's turn in the game state shown above. MAX's utility of wining, losing, and drawing is 1, -1, and 0, respectively.

b) Draw the complete game tree for MAX in this state. Assume in the tree that each player chooses the columns from left to right. Write the utility next to each terminal node. Turn the paper in "landscape mode" and simplify the representation of each state by only drawing the free slots as shown below.



c) Write the minimax value next to each node in the game tree that you made in your answer to sub-problem b) and indicate the minimax decision of MAX (you do not have to re-draw the tree).

d) Indicate the alpha and beta cuts that the MINIMAX algorithm with α-β pruning would have made in the tree that you made in your answer to sub-problem b) (again, you do not have to re-draw the tree). Is it possible to prune more nodes if MAX in the root node chooses columns in another order than left to right (why/why not)?

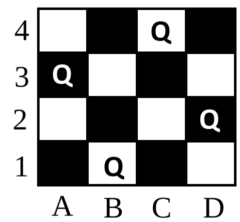
Since the game trees of Connect 4 are very large, the MINIMAX algorithm must apply an evaluation function and cutoff test to be computationally feasible. Consider two evaluation functions f_1 and f_2. They are both defined in terms of lines of connected discs with same color in a game state. Each line must contain at least two discs in the horizontal, vertical or diagonal direction and each line must have maximum length. As an example, the game state shown above has three vertical black lines of length three and one vertical black line of length two. The evaluation functions differ by the value they assign to a line of length n. For f_1, the value is $2n$, while it for f_2 is $n!$. Both functions evaluate a game state as the total value of black lines minus the total value of white lines.

e) Consider two agents playing Connect 4 against each other using the same MINIMAX algorithm but one applying evaluation function f_1 and the other f_2. Which of the agents do you think would win and why?

f) Positive values of f_1 and f_2 indicate that the game state is strongest for black. For that reason we expect terminal states won by black to have positive f_1 and f_2 values. Is that always the case (why/why not)?

Problem 2: Constraint Satisfaction (40%)

Recall that the 4-Queens Problem consists in placing four queens on a 4×4 chessboard so that no two queens can capture each other. The figure below shows a solution.



We will use four variables $\{A, B, C, D\}$ with domain $\{1, 2, 3, 4\}$ to represent the positions of the queens. Each variable is associated with a column of the chessboard. The value of a variable is the row number of a queen placed in the column. Hence, the variable assignment representing the solution shown above is $A = 3$, $B = 1$, $C = 4$, and $D = 2$.

a) Prof. Smart claims that this is a bad representation of the problem since it is incomplete. It is possible to place more than one queen in a column, but the representation does not allow that. Do you agree (why/why not)?

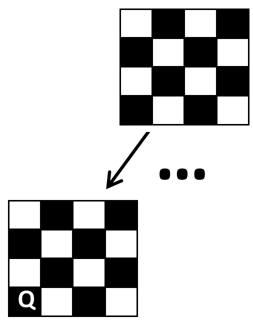
We will use a simple version of backtracking search called BASIC-BACKTRACK-SEARCH to find solutions to the 4-Queens Problem. The pseudocode of BASIC-BACKTRACK-SEARCH is shown below.

function BASIC-BACKTRACK-SEARCH(*csp*) **returns** a solution, or failure
  return BASIC-BACKTRACK($\{\}$,*csp*)

function BASIC-BACKTRACK(*assignment*,*csp*) **returns** a solution, or failure
  if *assignment* is complete **then return** *assignment*
  var $\leftarrow$ SELECT-UNASSIGNED-VARIABLE(*csp*)
  for each *value* **in** ORDER-DOMAIN-VALUES(*var*,*assignment*,*csp*) **do**
    if *value* is consistent with *assignment* **then**
      add $\{$ *var* $=$ *value* $\}$ to *assignment*
      result $\leftarrow$ BASIC-BACKTRACK(*assignment*,*csp*)
      if *result* $\neq$ *failure* **then**
        return *result*
    remove $\{$ *var* $=$ *value* $\}$ from *assignment*
  return *failure*

Assume that SELECT-UNASSIGNED-VARIABLE chooses the variables in the order A, B, C, D and that ORDER-DOMAIN-VALUES orders domain values $1, 2, 3, 4$.

b) Draw the search tree traversed by BASIC-BACKTRACK-SEARCH for the 4-Queens Problem. Each call to BASIC-BACKTRACK is a node in the tree and is drawn as the state of the chessboard given by assigned variables. The initial node and a child node corresponding to the assignment $A = 1$ are illustrated below. Mark failed leaf nodes with the letter "F" and mark the solution node with the letter "S".



BASIC-BACKTRACK-SEARCH returns a single solution or failure if no solution exists. We now would like to modify it to find all solutions.

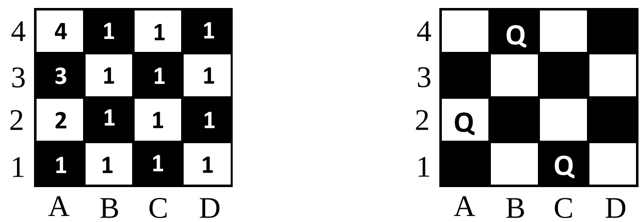
c) Assume that BASIC-BACKTRACK-SEARCH continues searching after finding the first solution to the 4-Queens Problem. Draw the complete search tree it then explores. Mark failed leaf nodes with the letter "F" and mark solution nodes with the letter "S".

Let BACKTRACK-SEARCH-ALL refer to this new version of BASIC-BACKTRACK-SEARCH.

d) Write pseudocode for the function BACKTRACK-SEARCH-ALL(*csp*) that returns the set of all solutions to a given constraint satisfaction problem, *csp*. Notice that the function simply returns the empty set if *csp* has no solutions. Hint: use the pseudocode for BASIC-BACK-TRACK-SEARCH as a template.

e) Can the *least constraining value heuristic* for ordering domain values be used in order to reduce the number of search nodes explored by BACKTRACK-SEARCH-ALL (why/why not)?

Assume that each domain value of the variables in a constraint satisfaction problem is associated with a cost. Further, define the cost of a variable assignment to be the sum of the costs of the assigned values. For our representation of the 4-Queens Problem, this means that each position on the chessboard has a cost. Assume that these costs are defined as shown in the left figure below. Thus, the cost of the assignment $A = 2$, $B = 4$, and $C = 1$ shown in the right figure is $2 + 1 + 1 = 4$.



An algorithm that returns all solutions like Backtrack-Search-All can be used to find a solution with minimum cost of its assignment.

f) Calculate the costs of the assignments of the nodes in the complete search tree of the 4-Queens Problem that you made in your answer to sub-problem c). Write the costs next to the nodes and mark nodes with minimum cost solutions with a star "*" (you do not have to re-draw the tree). Assume that Backtrack-Search-All finds a minimum cost solution by keeping track of a least cost solution found so far during the search. How can the cost of this solution be used to prune branches of the search tree that only can contain sub-optimal solutions? Show which branches are cut in the tree.

Problem 3: Logic (30%)

Patients often infect each other and it is important for hospitals to know whether an outbreak is COVID-19 or just flu or cold. We will develop a small knowledge base to determine this. Let the propositional symbols X, Y, and Z identify whether a group of patients have COVID-19, flu, or cold, respectively. Thus, if X is *true*, the group has COVID-19. The table below describes two relevant symptoms of these diseases.

Symptom	**COVID-19** (X)	**Flu** (Y)	**Cold** (Z)
Fever (F)	Common (*true*)	Common (*true*)	Rare (*false*)
Loss of Appetite (L)	Common (*true*)	Sometimes (*false*)	Common (*true*)

We use the propositional symbols F and L to denote the observed symptoms of the patient group. F denotes fever and is *true* if fever is common in the group and *false* if it is rare. L denotes loss of appetite and is *true* if it is common and *false* if it sometimes happens in the group. Assume that we know that all patients in the group have one and only one disease and that it is either COVID-19, flu, or cold. Moreover, assume that we can draw conclusions from the symptoms that are certain. For instance, if fever is rare in the patient group, we can conclude that they have cold. The table below shows three different sentences (1), (2), and (3) representing the fever symptom.

(1)	(2)	(3)
$(F \Rightarrow X \wedge Y \wedge \neg Z)$	$(F \Rightarrow X \vee Y) \wedge$ $(F \Rightarrow \neg Z)$	$(F \Rightarrow X \vee Y) \wedge$ $(F \Rightarrow \neg Z) \wedge$ $(\neg F \Rightarrow Z) \wedge$ $(\neg F \Rightarrow \neg X \wedge \neg Y)$

a) Explain why sentence (1) is incorrect and sentence (2) does not represent all information about the fever symptom according to our assumptions.

b) Prof. Smart claims that sentence (3) is logically equivalent to sentence (2) if the implication operator in each conjunct of (2) is substituted with a bi-implication operator. Nevertheless, he prefers sentence (3), since each of the four conjuncts of (3) is a definite clause. Do you agree about these claims (why/why not)?

c) Draw the BDD of (3) using variable order $F \prec X \prec Y \prec Z$.

d) Write a sentence of the loss of appetite symptom (L) in the same format as (3).

e) Write the conjunction of (3) and the sentence for loss of appetite you gave in your answer to sub-problem d) in conjunctive normal form (CNF). Write all intermediate results of your derivation.

f) Assume that a knowledge base (KB) contains the CNF sentence that you gave in your answer to sub-problem e). Show that $KB \wedge F \wedge L \models X$ using resolution. That is, show that the patients have COVID-19 if fever and loss of appetite is common.