Dr. Eick

Fundamentals of Artificial Intelligence

COSC 4368

Solution Sketches Midterm Exam

Wednesday, March 6, 2024



*Name:*

*Student id:*

1. A\* & Greedy Best-first Search (13 points):
2. Game Theory (7 points):
3. Reinforcement Learning (14 points):
4. SA and Backtracking (8 points):
5. Miscellaneous Questions (10 points):
6. Constraint Satisfaction Problems (9 points):

Σ (out of 61):

Number Grade:

The exam is “open books and notes”, but no computers and cell phones allowed; you have 75 minutes to complete the exam. Write all your answers on this document (you can use back sides!).

**1) Best-first Search and A\* [13]**

Consider the search space below, where *S* is the start node and *G1* and G2 satisfy the goal test. Arcs are labeled with the cost of traversing them and the estimated cost to a goal (the h function itself) is reported inside nodes.

For each of the following search strategies, indicate which goal state is reached (if any) and list, *in order*, all the states *popped off of the OPEN list*. When all else is equal, nodes should be removed from OPEN in alphabetical order.

##### a) Best-First-Search (using function h only) [2]

Goal state reached: G2 [1]

States popped off OPEN: S E G2 [1]

##### b) A\* (using f=g+h)[5]

Goal state reached: G2 [1]

States popped off OPEN: S E A D B G2 [4]

No partial credit except up to 1 point partial credit for “b states poppled”

9

1

7

2

1

5

2

9

2

5

8

5

1

4

7

Problem 1 continued

b) What is the main difference between A\* and greedy best first search? [2]

A\* additionally considers the cost of reaching a node n from the starting node (g(n)) whereas greedy best first search solely considers the cost of reaching a goal node from node n (h(n))

c) Assume you have 2 admissible heuristics h1(s) and h2(s) are given for a given seach problem. Is h3(s)=max(h1(s),h2(s)) also admissible? Would you prefer using h2 or using h3 in conjunction with A\*? Give reasons for your answers![4].

Yes, h3 is admissible. [1]If h1 and h2 always underestimate the “true” cost then the higher of the two will certainly underestimate the true cost as well; therefore, h3 is admissible. [3]

I will prefer h3, as h3 is always higher than h2 and therefore it provides a closer approximation of the true cost. ~~As a matter of fact, h3 dominates h2, which translates into equal or better efficiency of the search, as discussed on the bottom of page 106 of our textbook.~~

**2) Game Theory [7]**

What is the Nash Equilibrium[[1]](#footnote-1) for the following parallel game, whose payoff matrix is depicted below [5]? Player 1 has actions A, B, C and D whereas Player 2 has actions X, Y, and Z. What is the main property of a Nash Equilibrium? [2]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | A | B | C | D |
| X | 2,2 | 2,5 | 6,1 | 3,9 |
| Y | 3,3 | 9,2 | 2,4 | 1,7 |
| Z | 2,1 | 1,2 | 3,4 | 1,1 |

There is one Nash equilibrium (C,Z) 0-1.5 points if they do not give that answer

If one player changes her action her payoff will not increase [2]

3) Reinforcement Learning [14]

a) What is the goal of reinforcement learning? [2]

**Reinforcement learning** (**RL**) centers how an intelligent agent ought to take [actions](https://en.wikipedia.org/wiki/Action_selection) in a dynamic environment in order to maximize its [cumulative reward](https://en.wikipedia.org/wiki/Reward-based_selection). Or

Reinforcement Learning centers learning policies that enable an agent to act intelligent in a potentially unknown or changing world.

Other answers might deserve full or partial credit.

b) Consider the following World called ABC is given:

A diagram of a mathematical equation

Description automatically generated

Give the Bellman equations for states 2, 3, 4 for the ABC World, assuming γ=0.2. [4.5]

Ut+1(2)= -2 + 0.2\*Ut(4) [1]

Ut+1(3)= -9 + 0.2\*mas (Ut(4), Ut(2)) [1.5]

Ut+1(4)= 3 + 0.2\*max(0.4\*Ut(1)+0.6\*Ut(5), 0.1\*Ut(1)+0.6\*Ut(5)+0.3\*Ut(3) [2]

No partial creadit for formulas with errors. 0.5 points penalty is γ is not expanded.

c) What is the pupose of the hyper parameter γ in Bellman equations? What is the impact of choosing a low value of γ? [2.5]

It measures the future wellbeing of the agent or [The discount factor is a **parameter that determines how much future rewards are valued** in reinforcement learning](https://www.bing.com/ck/a?!&&p=9c82f477d11f7d42JmltdHM9MTcxMDgwNjQwMCZpZ3VpZD0xMTNhMTQ2Zi03YjRiLTZkNGYtMWU2ZC0wNDZhN2FhZDZjZmMmaW5zaWQ9NTg0Ng&ptn=3&ver=2&hsh=3&fclid=113a146f-7b4b-6d4f-1e6d-046a7aad6cfc&psq=discount+factor+reinforcement+learning&u=a1aHR0cHM6Ly9pbnR1aXRpdmV0dXRvcmlhbC5jb20vMjAyMC8xMS8xNS9kaXNjb3VudC1mYWN0b3Iv&ntb=1)[**1**](https://www.bing.com/ck/a?!&&p=46c46d24287961f8JmltdHM9MTcxMDgwNjQwMCZpZ3VpZD0xMTNhMTQ2Zi03YjRiLTZkNGYtMWU2ZC0wNDZhN2FhZDZjZmMmaW5zaWQ9NTg0Nw&ptn=3&ver=2&hsh=3&fclid=113a146f-7b4b-6d4f-1e6d-046a7aad6cfc&psq=discount+factor+reinforcement+learning&u=a1aHR0cHM6Ly9pbnR1aXRpdmV0dXRvcmlhbC5jb20vMjAyMC8xMS8xNS9kaXNjb3VudC1mYWN0b3Iv&ntb=1)[**2**](https://www.bing.com/ck/a?!&&p=49493b9473b3ff87JmltdHM9MTcxMDgwNjQwMCZpZ3VpZD0xMTNhMTQ2Zi03YjRiLTZkNGYtMWU2ZC0wNDZhN2FhZDZjZmMmaW5zaWQ9NTg0OA&ptn=3&ver=2&hsh=3&fclid=113a146f-7b4b-6d4f-1e6d-046a7aad6cfc&psq=discount+factor+reinforcement+learning&u=a1aHR0cHM6Ly9jcy5zdGFja2V4Y2hhbmdlLmNvbS9xdWVzdGlvbnMvNDQ5MDUvdGhlLW1lYW5pbmctb2YtZGlzY291bnQtZmFjdG9yLW9uLXJlaW5mb3JjZW1lbnQtbGVhcm5pbmc&ntb=1)

*Problem 3 continued*

d) We apply temporal difference learning to the DEF World depicted below, assuming the agent starts in state 2 and applies the operator sequence **sw-n-e**; assume that the initial utilities are 0; what are the new utilities[[2]](#footnote-2) of the 3 states have been visited by the agent? Also assume α=0.5 and γ=1! [5]

Diagram

Description automatically generated

**U(2) = 0.5 \* U(2) + 0.5 (R(2) + 1 \* U(4)) =** 7

**U(4) = 0.5 \* U(4) + 0.5 (R(4) + 1 \* U(1)) =** -2

**U(1) = 0.5 \* U(1) + 0.5 (R(1) + 1 \* U(2)) =** -1.5

If one error: at most 2 points

If more than 1 error: at most 1 point

**4) Simulated Annealing and Backtracking [8]**

a) Simulated annealing allows taking downward moves with respect to the objective function; what does the probability of allowing downward steps depend on? [3]

a) to which extend the run of SA is complete / the temperature [1.5]

b) the extend of the downward step / how much was the new solution is in comparison to the current solution [1.5]

Problem 4 continued

b) Assume you apply a version of backtracking that checks for repeated states on the current path[[3]](#footnote-3), but which does not use a depth bound to the 8-puzzle. Will it always find a solution if a solution exists (assuming that there are enough computational resources)? Give reasons for your answer! [5]

Yes! [1]

Because the search-tree for the 8-puzzle is finite and because the algorithm checks for loops in the current path the algorithm will sooner or later stop moving forward, backtrack, and find the solution eventually. [4]

**5) Miscellaneous Questions (10 points)**

a) Assume you use alpha-beta search for a game such as Go or chess where generating the complete game tree is not feasible as it is too large. How do you deal with this problem? [3]

Create a state evaluation function which evaluates the goodness of the state for non-terminal states [1.5] and limit the search by using a depth bound. [1.5]

0.5 points penalty if they do not mention nonterminal states.

b) How do mutation operators do in evolutionary computing systems? What role does mutation play with respect to the gene pool in the current population? [3]

It performs random changes to a solution. [2] Mutation is a exploration operator that introduces new values in the mutated solution. [1]

c) Evolutionary computing systems are based on the Darwinian Evolution and survival of the fittest. How do they accomplish that? How is Darwinian Evolution simulated in evolutionary computing systems? [4]

a. They employ selection methods that enable fitter individual to reproduce with a higher probability. [2] b. E.g. by assigning members of the population a portion of a roulette wheel that is proportional to their fitness making it more likely that fitter individuals are chosen to reproduce.

Other answers for b might deserve full are partial credit.

6. **Discrete CSPs (9 points)**

Assume the following constraint satisfaction problem in which variables A, B, C, D take values in {1,…,100}is given:

* **Constraints:**
  + (C1) A < B
  + (C2) C\*D + B\*D\*D=D\*D\*D\*A
  + (C3) B\*B\*C + A\*B\*D=A\*A\*D\*D
  + **(**C4) C < D

A brute force solution to this problem could look as follows:

FOR A=1,…,A=100

FOR B=1,…,B=100

FOR C=1,…,C=100

FOR D=1,…,D=100 DO {

IF C1 and C2 and C3 and C4 THEN WriteSolution(A,B,C,D)}

Give the code of a more efficient solution to this problem which uses less loops and/or less iterations inside the loop. Briefly describe the idea of your solution!

a. simplify C2 by dividing by D to: (C2’) C=D\*D\*A-B\*D

b. Eliminate C and take advantage of C1:

FOR A=1,…,A=100

FOR B=A+1,…,B=100

FOR D=1,…,D=100 DO {

C= C=D\*D\*A-B\*D

If C3 and C4 THEN WriteSolution(A,B,C,D)} 9 points

Alternative Solution which takes advantage of C1 and C4

FOR A=1,…,A=100

FOR B=A+1,…,B=100

FOR C=1,…,C=100

FOR D=C+1,…,D=100 DO {

IF C2 and C3 THEN WriteSolution(A,B,C,D)} only 7 points

Other solutions might deserve full or partial credit!

1. There could be also none or more than 1! [↑](#footnote-ref-1)
2. Use  [↑](#footnote-ref-2)
3. This version backtracks, if a loop in the current path is encountered. [↑](#footnote-ref-3)