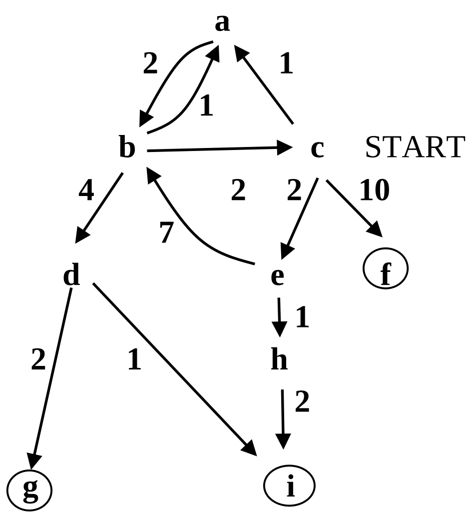
Artificial Intelligence, Master degree - Exam 11 June 2020

Instructions:
Time: 70 min. Use of literature, notes and electronic devices is not allowed. Please state your answers short and clear, answering the questions directly to the point.

1. Consider the following state space:



Let **c** be the start state of search. **f**, **g** and **i** are goal states. Let search algorithms generate the successor nodes of a node in alphabetical order. For example, the order of successors of node **c** is: **a**, **e**, **f**.

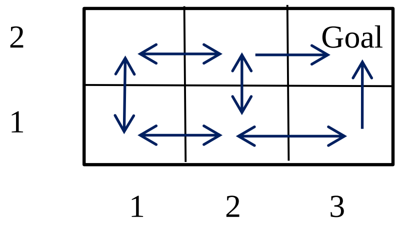
Assume that search algorithms A*, IDA* and RBFS detect cycles, and they immediately reject a node that completes a cycle. But they handle a graph as a tree. That is, if the algorithm reaches a node N by an alternative path then a copy N' of N is created and N' is treated as a new node. If two nodes have equal f-value then the node that was generated first is expanded first.

Heuristic values h of the nodes are given in the following table:

X	a	b	c	d	e	f	g	h	i
h(X)	1	1	1	2	4	0	0	8	0

(a) Which solution path is returned by algorithm A*? Give the order nodes in which the nodes are expanded during search by A*.
(b) Which solution path is returned by algorithm IDA*?
(c) Which solution path is returned by algorithm RBFS?
(d) During the execution of RBFS in this case, what is the value of bound B for searching the subtree below node e? How is the backed-up value F(a) changing?
(e) Are there any nodes whose F-values are *inherited* from their parents during the execution of RBFS in this case? If yes, state these nodes and their inherited values.
(f) Which solution path is returned by the RTA* algorithm if look-ahead depth is 2 steps? What is the stored h-value of node c?

2. The 2-dimensional grid 3x2 below specifies a reinforcement learning problem.



A mobile robot can move in the grid between the cells. The arrows indicate possible robot's actions and corresponding transitions between the states. The actions are l, r, d, u (left, right, down, up). As shown in the figure above, not all actions are possible in every state. In state s32 (i.e. topmost rightmost cell (3,2)), no action is possible (goal state). The system is deterministic except for the action »right« in states s11 and s22. Here the action »r« causes the intended transition to the right with probability S, but it may also happen due to mechanical problems that the robot stays where it was before the action (with probability 1-S). So for example after action "r" in state s22, the next state may be s32 or s22. Assume S = 0.2. Rewards for all the transitions are equal 0, except for the transitions into the goal state s32. Transitions s22-s32 and s31-s32 are rewarded by 100. The cumulative reward is discounted by factor gamma = 0.5.

(a) What is the cumulative reward after the action sequence [d,r,u] from the state s22?

(b) What is the expected immediate reward R(s22,r) if action "r" executed in state s22?

(c) What is the utility U^r_{22} of state s22 as the expected cumulative reward of a policy that always executes action "r" in s22 ?

(d) Starting at s11, there are two reasonable candidate policies: (1) "Low path": try to go to the goal following the path s11, s21, s31, s32, or (2) "High path": try to go to the goal following the path s11, s12, s22, s32. Which of these two policies, high path or low path, has a higher utility for state s11? Justify your answer.

(e) Now suppose that S is not given. The best action in s22 depends on S. The best action might be "r" or "d" (followed by "r", "u"), depending on S. For high S, action r is better, for low S action d is better. What is the critical value of S when the choice between the two policies changes?

3.

(a) How is a final plan represented in the POP approach to planning?

(b) How is a plan represented in the GRAPHPLAN approach to planning?

(c) In the GRAPHPLAN method, there are three types of mutual exclusion (mutex) between two actions. How is the "interference" type of mutex between actions **A1** and **A2** defined in terms of preconditions and effects of **A1** and **A2**? Use the following notational convention: **can(A)** denotes the set of preconditions of action **A**, and **effects(A)** denotes the set of effects (both positive or negative) of **A**.

(d) Describe the time complexity of the GRAPHPLAN algorithm in terms of the number of literals and actions in the planning graph. Specify the order of time complexity separately for (1) the construction of the planning graph, and (2) the extraction of a plan from the planning graph.

4. Consider a qualitative model of QSIM type. There are 4 variables in the system: X, Y, VX, VY. The landmarks for these variables are:

X: minf, a, 0, b, inf
Y: minf, a, 0, b, inf
VX, VY: minf, 0, inf

The constraints in the model are:

```
deriv( X, VX)
deriv( Y, VY)
plus( X, VX, b/std)     % VX = b/std - X  (VX is difference between landmark b and X)
plus( Y, VY, X),        % VY = X - Y  (VY is diff. between X and Y)
```

$VY = M_0^+(VX)$ % Monotonically inc. function with corresponding values (zero,zero)

Initial values of X and Y in the start state at time t0 are: X(t0) = 0, Y(t0) = a

(a) Determine the qualitative state of the system at time t0; that is qualitative magnitudes and directions of change of all the variables: X, Y, VX, VY.

(b) Determine the qualitative state of the system at time interval t0..t1.

(c) Determine the qualitative state of the system at time t1.

(d) If we continue the simulation, the system eventually reaches a steady state. What is this state? Describe intuitively the mechanics of how this system works.