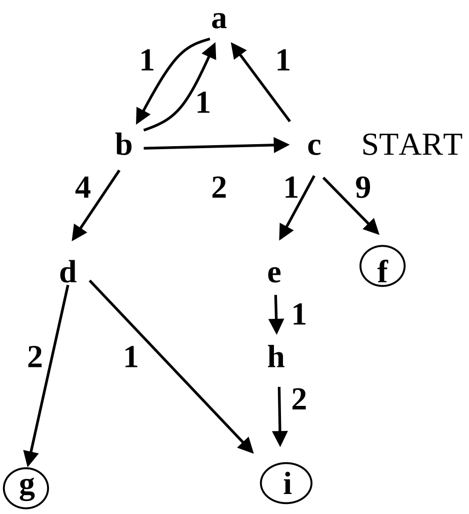
**Artificial Intelligence, Master degree - Exam 20 June 2019**

**Instructions:**
Time: 75 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.
Oral exam: Monday, 24 July 2019 at 14 pm

**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 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 | 2 | 1 | 2 | 2 | 0 | 0 | 8 | 0 |

(a) Which solution path is returned by algorithm A*?
(b) Which solution path is returned by algorithm IDA*? How is the f-limit changing during the execution of IDA* in this case?
(c) Is evaluation function f in this search problem monotonic? Briefly justify your answer.
(d) Which solution path is returned by algorithm RBFS?
(e) In the execution of RBFS, what is the value of bound B for searching the subtree below node a?
(f) Give the sequence of current states during the execution of RTA* if lookahead depth is 1? What are the stored h-values of nodes c and a?

**2.**

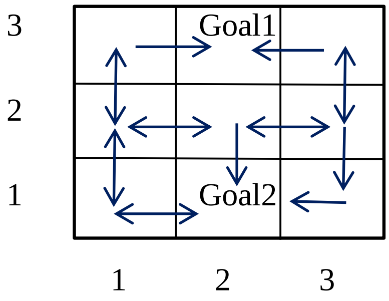
(a) The time complexity of the GRAPHPLAN planning method roughly depends on the length D of a solution plan, and on two other parameters of the planning problem. Explain what are these two other parameters?

(b) The GRAPHPLAN planning method consists of two steps: (1) construction of the planning graph, and (2) extraction of a plan from the planning graph. Describe roughly the order of time complexity of each of these two steps in terms of the parameters from question (a).

(c) Let the notation **can(A)** denote the set of preconditions of an action **A**, and **effects(A)** denote the set of effects (both positive or negative) of **A**. Let **A1** and **A2** be two actions. One type of mutex relation between actions is "interference". State the necessary and sufficient condition (in terms of **can(A1)**, **can(A2)**, **effects(A1)** and **effects(A2)**) for no interference between **A1** and **A2**.

(d) What are two main uses of planning graphs in construction of plans?

**3**. Consider the following reinforcement learning domain where an agent is moving on a 3x3 grid:



The arrows show what are possible actions in each state. The system is deterministic except for the action »left« in the state (3,1) (bottom rightmost corner). In this case, the result of action »left« can be actually moving to the left with probability 0.6, and moving upwards with probability 0.4. There is no action possible in the goal states (2,3) and (2,1). Rewards for all the state transitions are 0, except for the following:

Transition to any of the goal states gives reward 1  
Transition from (3,2) to (3,1) gives reward 2

The cumulative reward is discounted by factor gamma = 0.5.  
Let the start state be (2,2).

Let policy PI be defined as follows:  
PI ((2,2)) = right, PI((3,2)) = down, PI((3,1)) = left

(a) What is the utility $U^{PI}((3,1))$ of state (3,1) given this policy? Use the abbreviated notation $U_{31} = U^{PI}((3,1))$.

(b) What is the utility $U_{22}$ of state (2,2) given policy PI? ($U_{22} = U^{PI}((2,2))$ )

(c) Suppose that a TD-learning agent (who does not know a model of the above system) is trying to determine the utilities of states (2,2), (3,2), and (3,1) for the policy PI. Let initial approximations of these utilities be: $U_{22} = U_{32} = U_{31} = 0$, and parameter alpha of theTD-update rule be alpha = 0.8. Suppose the agent has conducted two trials. Trial 1 resulted in the sequence of observed transitions between states: (2,2) --> (3,2) --> (3,1) -->(2,1). What are the approximations of the utilities $U_{22}$ , $U_{32}$ , and $U_{31}$ after Trial 1?

(d) Suppose Trial 2 also resulted in the same transitions as Trial 1. What is the approximations of the utilities $U_{22}$ , $U_{32}$ , and $U_{31}$ after Trial 2?

**4.** Consider a qualitative model of QSIM type. There are 5 variables in the system: X, Y, D, VX, VY. The landmarks for these variables are:

X: minf, x0, 0, inf
Y: minf, 0, y0, inf
D, VX, VY: minf, 0, inf

The constraints in the model are:

```
deriv( X, VX)
deriv( Y, VY)
plus( X, D, Y)     % D = Y – X
VX = M0+(D)        % Monotonically inc. function with corresponding values (zero,zero)
VY = - VX          % Opposite velocities VX and VY
```

Initial values of X and Y in the start state at time t0 are: X(t0) = x0, Y(t0) = y0

(a) Determine all possible qualitative states of the system at time t0; that is qualitative magnitudes and directions of change of all the variables: X, Y, D, VX, VY.
(b) Determine all possible states of the system at time interval t0..t1.
(c) Determine all possible qualitative states of the system at time t1.
(d) Now consider the same system, but without the constraint: $VX = M_0^+(D)$. Let the values of X and Y at time t0 be again: X(t0) = x0, Y(t0) = y0. How many qualitative states of the system are possible now at time t0?