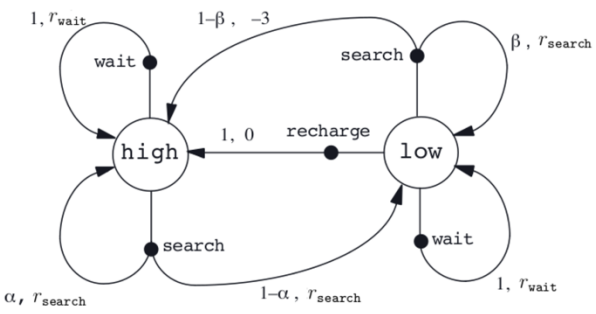
POLITECNICO MILANO 1863	**AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS** September 8, 2021		LAST NAME AND FIRST NAME
	ROW	COLUMN	ID NUMBER (CODICE PERSONA)

- The exam is composed of three stapled sheets printed on both sides.
- This front page must be filled with last name, first name, ID number, position (row and column communicated by the instructor), and signature.
- Exams without a completely filled front page or with missing sheets will not be considered.
- Answers can be written only on these sheets. If you need more space, please write on the last page.
- Exam is open books.
- **All the answers must be justified.**

SIGNATURE

Question 1. Consider the Markov Decision Process (MDP) reported in the figure, in which the values (parametric, in some cases) for the transition function and for the reward function are reported on the arcs.



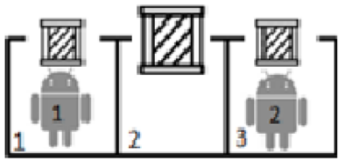
Answer the following questions, clearly motivating your answers (e.g., reporting calculations).

(1) (3 points) Report the Bellman equations for the two states (low and high). Please be as specific as possible, according to the information available in the figure.
(2) (2 points) Assuming $\alpha = 1$ and a discount factor equal to 0.8, is it possible that the optimal policy performs action wait when the agent is in state high? If yes, report under which conditions (inequalities) it happens. If not, explain why.
(3) (2 points) Assuming all the unspecified quantities in the figure as parameters, is there any value of the parameters such that the optimal policy in state low prescribes to perform action recharge? If yes, report the values (no need of reporting exact values: an intuition is enough). If not, explain why.

Question 2. Consider a generic super-additive coalition game $G = (A, v)$ and answer the following questions, justifying the answers (e.g., using examples and counterexamples).

(1) (2 points) Can G have an empty core?
(2) (1.5 points) Can the Shapley value of an agent of G be 0?
(3) (1.5 points) Can the Shapley value of an agent of G be equal to the value of the grand coalition?
(4) (1.5 points) When applying the dynamic programming algorithm for finding the coalition structure, can the algorithm return a partition of A in 2 coalitions?
(5) (1.5 points) When applying the Shehory and Kraus algorithm for finding the coalition structure, can the algorithm return a partition of A in 2 coalitions?

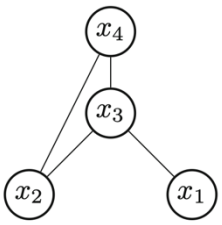
Question 3. Consider the box pushing domain in the figure. There are 3 rooms and 2 robots, which start in rooms 1 and 3. The robots can move from a room to an adjacent room and push a box to a warehouse (not shown in the figure; after pushing a box to the warehouse, a robot automatically returns to the room where the box initially was). The left and right boxes are light and can be pushed by a single robot, while the middle box is heavy and requires the two robots pushing together. The goal is to have all boxes in the warehouse.



Answer the following questions, motivating your answers.

(1) (3 points) Formalize the above setting as a Dec-POMDP and precisely define all the composing elements. Assume a finite horizon, that all the actions are deterministic, and that the observations are perfect.
(2) (2 points) Report the first 3 layers (including the root) of a policy tree for the left agent.
(3) (2 points) Given the formalization of (1), how is a joint policy evaluated?
(4) (1 point) Given the formalization of (1) what should an optimal joint policy do (informally)? Why?

Question 4. Consider the Distributed Constraint Optimization Problem (DCOP) shown in the figure. Values of the constraints are utilities.



(a)

x_1 \ x_3	0	1
0	1	7
1	9	3

x_2 \ x_3	0	1
0	7	3
1	8	3

x_3 \ x_4	0	1
0	4	7
1	6	6

x_2 \ x_4	0	1
0	3	3
1	1	5

(b)

Assume that the DCOP is solved using the Max-Sum algorithm.
(1) (2 points) What are the messages sent at the beginning of the algorithm by agent x_2? (Ignore the normalization factors α.)
(2) (2 points) Assuming that computations and communications are synchronous, what are the messages sent by agent x_1 at the second step of the algorithm? (Also in this case, ignore the normalization factors α.)

Assume that the DCOP is solved using the DSA algorithm and that the initial values chosen by agents are all 0 and that such values have been already communicated by the agents.
(3) (2 points) What are the next message sent by agent x_2 (assuming that it is activated)?

Assume that the DCOP is solved using the MGM algorithm and that the initial values chosen by agents are all 0 and that such values have been already communicated by the agents.
(4) (2 points) What is the next message sent by agent x_2?