Chair of Robotics, Artificial Intelligence and Real-time Systems
Department of Informatics
Technical University of Munich

Esolution
Place student sticker here

Note:

- During the attendance check a sticker containing a unique code will be put on this exam.
- This code contains a unique number that associates this exam with your registration number.
- This number is printed both next to the code and to the signature field in the attendance check list.

Grundlagen der künstlichen Intelligenz

Exam:	IN2062 / Mock Exam	**Date:**	Saturday 1st January, 2022
Examiner:	Prof. Dr.-Ing. Matthias Althoff	**Time:**	14:00 – 15:30

Working instructions

- This exam consists of **16 pages** with a total of **9 problems**.
  Please make sure now that you received a complete copy of the exam.
- The total amount of achievable credits in this exam is 52.5 credits.
- Detaching pages from the exam is prohibited.
- Allowed resources:
  - a **pen or PDF editor** (do not write with red or green colors nor use pencils)
  - a **non-programmable pocket calculator**
  - **2 pages summary** (**1 double-sided** A4 page), **handwritten**
  - **empty scratch paper (do not submit)**
- Please write answers on the exam booklet only. If you run out of space, write on the additional pages provided. Notes on other paper will be disregarded.
- You must hand in all pages of the exam.
- **Answers are only accepted if the solution approach is documented.** Give a reason for each answer unless explicitly stated otherwise in the respective subproblem.
- All subproblems are solvable independently from each other if not explicitly stated differently.
- **Multiple-Choice questions are evaluated automatically.** Use a cross to select your answer:

  ☐ Answer A

  ☒ Answer B

  If you want to correct your answer, fill out the checkbox, and cross your new answer:

  ☒ Answer A

  ■ Answer B

  Notes next to the checkboxes cannot be evaluated.

Left room from ________ to ________ / Early submission at ________

Problem 1 Search (8.5 credits)

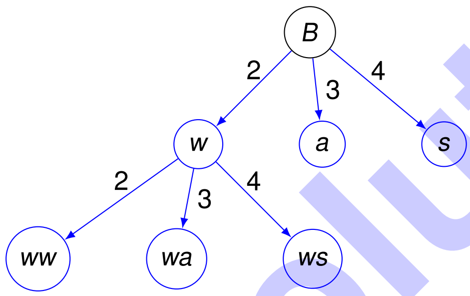
We want to build a bridge. At each step, you can use one piece of any material listed below to increase the length of the bridge, which also increases its weight. The properties of the 3 available materials are given in the table below. The stock of material is unlimited.

piece of material	increase in length	increase in weight
wood (w)	2 m	4 kg
aluminum (a)	3 m	7 kg
steel (s)	4 m	10 kg

We model this problem as a search problem. Assume that we start at node B, the base on one of the bridge's ends. We name expanded nodes according to the materials used to reach this node; for example, after using one piece of wood (w)in the first step and a piece of steel (s)in the second step, the node is labeled as *ws*.

0 1 2

a) Draw the search tree/graph with all possible nodes until each leaf node represents a bridge with a length of at least 3 m. Next to each arc, annotate the path cost (from node to node) for **length**. Do not draw more nodes than necessary.



+1 for correct nodes with labels (also if *wa* and *ws* are missing).
+1 for correct path costs (-0.5 for each wrong or if all based on weight). -0.5 if less than four nodes.

0 1

b) Your task is to build a 11 m long bridge with as few pieces as possible. What **uninformed** search method should you use? What is a possible goal node?

Search method: Breadth-First (or Iterative Deepening or UCS with path costs proportional to steps) +0.5

Goal node: ___ *ssa* or *sas* or *ass* ___ +1 for any of these

0 1 2

c) Your task is to build a bridge which is as lightweight as possible. What **uninformed** search method and edge costs should you use? After exploring node *B*, what are the next four nodes in the order they are explored?

Search method + edge costs: Uniform-Cost (or Dijkstra) +0.5 with mass of piece as cost +0.5

Explored nodes: *B*, ___ *w* ___, ___ *a* ___, ___ *w w* ___, ___ *s* ___.
1 point; -0.5 for each wrong or switched with neighbor

To avoid page flipping, we print the table again:

piece of material	increase in length	increase in weight
wood (w)	2 m	4 kg
aluminum (a)	3 m	7 kg
steel (s)	4 m	10 kg

d) Your task is to build a bridge with the weight (in kg) as path costs and the following heuristic function: $h(n) = 20 - \text{num}(w, n) - 9 \cdot \text{num}(a, n) - 6 \cdot \text{num}(s, n)$, where $\text{num}(m, n)$ is the number of pieces of material m used up to node n; e.g. the node *wa* has a heuristic value of $20 - 1 - 9 \cdot 1 - 6 \cdot 0 = 10$. Children of a node are created in the order: wood (w), aluminum (a), steel (s).

1) Perform **Greedy-Best-First Graph-search** until the first created child node has an evaluation value $f(n) \leq 0$. Document your search in the table below by listing the nodes in the order they are created. Note that we provided additional rows in the table and you do not have to fill all.

2) What is the path cost of the first node with $f(n) \leq 0$?

0 1 2 3

1)

Node name	Evaluation function $f(n)$	Parent name
B	20	None
w	19	B
a	11	B
s	14	B
aw	10	a
aa	2	a
as	5	a
aaw	1	aa
aaa	-7	aa

+0.5 for each correct level with correct values, + 0.5 if whole search tree is correct and in right order.

2) Path cost to reach first node with $f(n) \leq 0$: 21 kg +1 for correct value

Problem 2 Searching Agents (3 credits)

The following task uses the grid world shown on the right. It consists of tiles denoted as in chess by two-dimensional coordinates (columns A to F and rows 1 to 6). Each tile has a number associated with it. The agent starts at A1 and follows this program:

	A	**B**	**C**	**D**	**E**	**F**
1	6	4	6	4	3	4
2	5	4	3	2	1	3
3	6	5	4	4	3	5
4	5	6	3	7	8	9
5	7	5	4	2	3	8
6	6	6	5	3	1	0

- The agent perceives the numbers in its 8 neighborhood[1]. Cells outside the shown grid are perceived as infinity.
- In each step, the agent moves to the cell with the lowest number in its 8 neighborhood. If there are multiple it chooses the first tieing cell in clockwise order starting with the cell to its right. E.g., in cell A4 it would choose B5 over B3.

0
1

a) State the kind of agent implemented. Briefly justify your answer.

Agent type: reflex agent
Reason: agent program only depends on current perception of the environment.

0
1

b) State the next four fields visited by the agent after A1. Would it eventually reach F6?

Steps: A1 - B1 - C2 - D2 - E2
Reaches F6: No

[1] Cells above/below, left/right, and diagonal from its current position

Problem 3 Solving a Constraint Satisfaction Problem (CSP) (5 credits)

Consider the constraint graph of a Constraint Satisfaction Problem (CSP) with four variables given in Fig. 3.1.

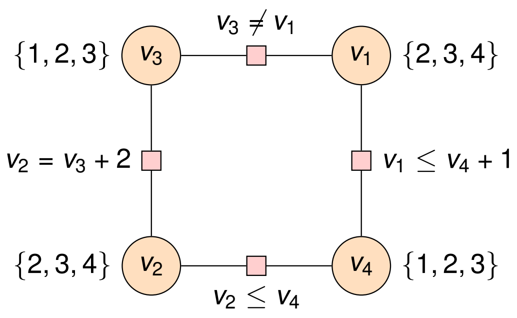


Figure 3.1: Constraint graph

The domains of and constraints on each variable are shown in Fig. 3.1.

a) Is the graph in Fig. 3.1 arc-consistent? If no, which arc(s) is/are not arc-consistent?

☐ Yes.

☐ No. ■ +0.5 for "No"

Arc(s):

(v_3, v_2) and (v_2, v_4) also "reverse" of these arcs is correct

+1 for each correct arc, -1 for each wrong arc

b) Assume that we **assign** $v_4 = 2$. Perform **forward checking** for the graph in Fig. 3.1.
1) What are the resulting domains of the other variables (not v_4)?
2) Based on your result, would you undo the assignment, i.e., backtrack? Why?

1) $v_3 \in \{1,2,3\}$ +0.5 $v_1 \in \{2,3\}$ +0.5 $v_2 \in \{2\}$ +0.5

2)

☐ Yes.

☐ No. ■

Reason:

you should not backtrack, since no domain is empty

+1 for correct decision (based on student's answer) incl. valid reason, -0.5 if valid reason is missing

Problem 4 **Propositional Logic (6 credits)**

Bob is a preschool teacher in Garching, preparing lunch for the children. He has certain ingredients and utensils at his disposal that he can choose to use or not. These are symbolized by the following propositional variables:

S : *Salt*	M : *Meat*
P : *Pan*	V : *Vegetables*
O : *Oven*	F : *Fruits*

0 1 2 3

a) Formulate Bob's cooking knowledge using propositional logic:

1. Salt has to be added to vegetables or meat, but not to fruits.

$[(V \lor M) \Rightarrow S] \land [F \Rightarrow \neg S]$

2. If vegetables and an oven are used, fruits or meat cannot be used.

$(V \land O) \Rightarrow \neg(F \lor M)$

3. A pan or an oven have to be used if and only if meat is to be cooked.

$(P \lor O) \Leftrightarrow M$

0 1 2 3

b) Bob also has 3 different types of spices, symbolized as A, B and C. He uses the following rule to determine which one to use:

$$\neg A \Leftrightarrow (\neg B \land \neg C)$$

Write this rule in conjunctive normal form.

$$\neg A \Leftrightarrow (\neg B \land \neg C)$$
$$[\neg A \Rightarrow (\neg B \land \neg C)] \land [(\neg B \land \neg C) \Rightarrow \neg A]$$
$$[A \lor (\neg B \land \neg C)] \land [\neg(\neg B \land \neg C) \lor \neg A]$$
$$[(A \lor \neg B) \land (A \lor \neg C)] \land [B \lor C \lor \neg A]$$
$$(A \lor \neg B) \land (A \lor \neg C) \land (B \lor C \lor \neg A)$$

Problem 5 First-Order Logic (7.5 credits)

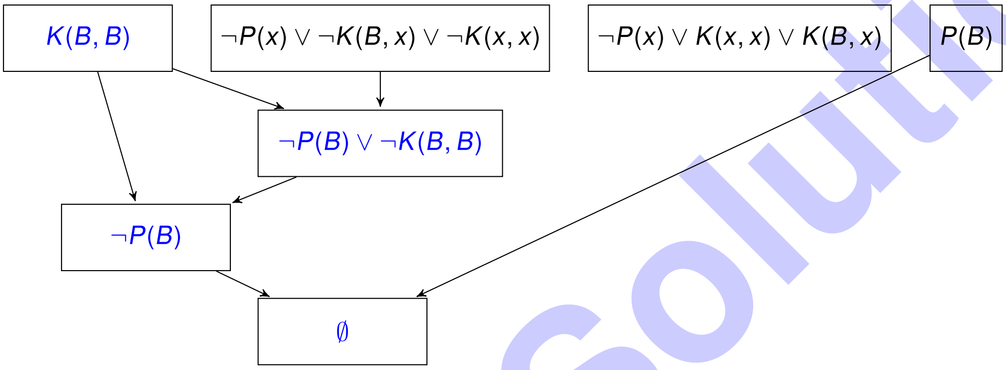
Barbara teaches the children in a kindergarten in Garching how to recognize their own names. Some children already know their own names, while others do not. This situation can be described using the predicates

$$P(x): x \text{ is currently in the preschool}$$
$$K(x,y): x \text{ knows the name of } y$$

and the constant B for Barbara. She decides to only remember the names of those who do not know their own names, a rule that can be expressed using the following knowledge base in conjunctive normal form:

$$\neg P(x) \vee \neg K(B,x) \vee \neg K(x,x)$$
$$\neg P(x) \vee K(x,x) \vee K(B,x)$$
$$P(B)$$

a) Complete the following resolution graph to show that the knowledge base entails $\neg K(B,B)$.



b) Using a similar argument as for a), one can show that the knowledge base also entails $K(B,B)$, so that the knowledge base entails both $K(B,B)$ and $\neg K(B,B)$. What is the meaning of $K(B,B)$ and $\neg K(B,B)$ in natural language? What can you deduce about the knowledge base?

The interpretation of $K(B,B)$ is that Barbara knows her own name. $\neg K(B,B)$ means that Barbara does not know her own name.
Since both can be entailed from the knowledge base, this means that the knowledge base is inconsistent/unsatisfiable.

c) Consider the case where only two people are involved: Barbara, symbolized by B, and Alice, symbolized by A. Transform the first-order logic sentence

$$\forall x, \quad K(B,x) \Leftrightarrow \neg K(x,x)$$

to a sentence in propositional logic without quantifiers. It does **not** need to be in conjunctive normal form, and you do **not** need to explain your result.

$$[K(B,B) \Leftrightarrow \neg K(B,B)] \wedge [K(B,A) \Leftrightarrow \neg K(A,A)]$$

d) Would you say that the sentence in c) is valid, satisfiable or unsatisfiable? Briefly explain your reasoning.

It is unsatisfiable. For example, $x = B$ falsifies the sentence, independently of the assignment of K. This can also be seen from c), since the sentence from c) is unsatisfiable.

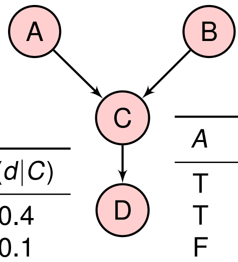
Problem 6 Bayesian Networks (4 credits)

0
1

a) Consider the Bayesian network below with three Boolean random variables.
Note: A = `True` is written as "a" and A = `False` is written as "$\neg a$". The same notation is used for all random variables. Compute the probability for A = `True` given that C = `False`.

P(a)
0.4

P(b)
0.2



C	**P**($d \mid C$)
T	0.4
F	0.1

A	B	**P**($c \mid A, B$)
T	T	0.6
T	F	0.5
F	T	0.2
F	F	0.1

$$P(a|\neg c) = \alpha P(a) \sum_b P(\neg c|a, B)P(B) \tag{6.1}$$
$$P(a, \neg c) = P(a)(P(\neg c|a, b)P(b) + P(\neg c|a, \neg b)P(\neg b)) \tag{6.2}$$
$$P(a, \neg c) = 0.4 \cdot (0.4 \cdot 0.2 + 0.5 \cdot 0.8) = 0.192 \tag{6.3}$$
$$P(\neg a, \neg c) = P(\neg a)(P(\neg c|\neg a, b)P(b) + P(\neg c|\neg a, \neg b)P(\neg b)) \tag{6.4}$$
$$P(\neg a, \neg c) = 0.6 \cdot (0.8 \cdot 0.2 + 0.9 \cdot 0.8) = 0.528 \tag{6.5}$$
$$\alpha = 1/(0.192 + 0.528) = 1.389 \tag{6.6}$$
$$P(a|\neg c) = 1.389 \cdot 0.192 = 0.27 \tag{6.7}$$
(6.8)

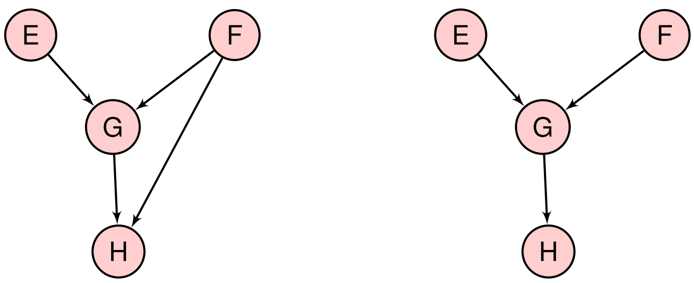
0
1

b) Would the probability computed in problem a) change if the **additional** information D = `True` was given? Only give the reason in text and don't perform any computations.

No, because A is conditionally independent of D given C.

0
1

c) Consider the structure of two Bayesian networks I and II sharing the same variables as shown below. Provide the condition that must be satisfied so that network I can be simplified to network II.



(a) Network I. (b) Network II.

Depending on randomized graph, either
$P(H|G, F) = P(H|G)$
or "H must be conditionally independent of F given G".

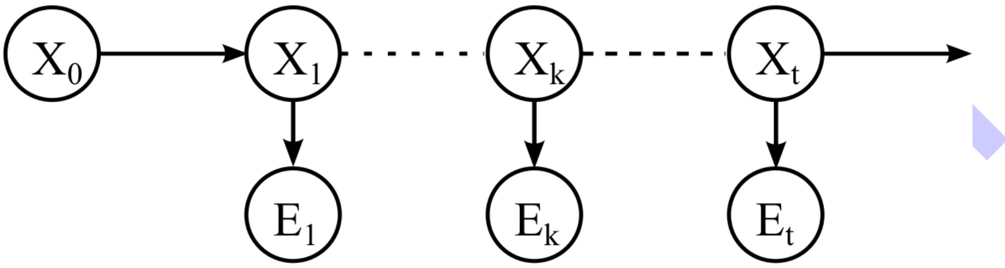
d) Suppose the probabilities $\mathbf{P}(U, w, x)$ shall be inferred from a Bayesian network with Boolean variables using enumeration. Below are four possible formulas given. Select the only one which can be correct **and** requires the smallest number of operations, i.e., multiplications and summations. (1 point)

- [ ] $P(x)P(w)\sum_{r} P(Y|R)P(R)\sum_{y} P(U|w, Y)$
- [ ] $P(x)\sum_{y} P(w)P(U|w, Y)\sum_{r} P(Y|R)P(R)$
- [x] $P(x)P(w)\sum_{y} P(U|w, Y)\sum_{r} P(Y|R)P(R)$
- [ ] $P(x)P(w)P(R)\sum_{y} P(U|w, Y)\sum_{r} P(Y|R)$

Problem 7 Hidden Markov Model (7 credits)

$P(X_0 = true)$
0.8

X_{k-1}	$\mathbf{P}(X_k = true \mid X_{k-1})$
t	0.7
f	0.3



X_k	$\mathbf{P}(E_k = a \mid X_k)$	$\mathbf{P}(E_k = b \mid X_k)$	$\mathbf{P}(E_k = c \mid X_k)$
t	0.5		0.3
f	0.5	0.3	

Consider the Hidden Markov Model above, where X_k is a Boolean random variable and E_k is a discrete random variable with domain $\{a, b, c\}$. Suppose that the evidence for $k = 1$ is $E_1 = c$, and for $k = 2$, the evidence is $E_2 = a$.

a) Fill in the missing values ($P(E_k = b|X_k = true)$ and $P(E_k = c|X_k = false)$) in the table of evidence. (The missing values can be written either in the gray boxes in the table above or in the solution box below.)

$P(E_k = b|X_k = t) = 0.2$
$P(E_k = c|X_k = f) = 0.2$

b) Calculate $P(X_2|E_{1:2})$.

$$f_{1:k} = P(X_k|E_{1:k})$$

$$f_0 = \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix}$$

$$T = \begin{bmatrix} P(X_k|X_{k-1}) & P(X_k|\neg X_{k-1}) \\ P(\neg X_k|X_{k-1}) & P(\neg X_k|\neg X_{k-1})) \end{bmatrix} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

Continuation of subproblem *b)*.

$$O_1 = \begin{bmatrix} P(E_1|X_1) & 0 \\ 0 & P(E_1|\neg X_1) \end{bmatrix} = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.2 \end{bmatrix}$$
$$O_2 = \begin{bmatrix} P(E_2|X_2) & 0 \\ 0 & P(E_2|\neg X_2) \end{bmatrix} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$

$$\begin{aligned} f_{1:1} = P(X_1|E_{1:1}) &= \alpha \cdot O_1 \cdot T \cdot f_0 \\ &= \alpha \cdot \begin{bmatrix} 0.3 & 0 \\ 0 & 0.2 \end{bmatrix} \cdot \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.7099 \\ 0.2901 \end{bmatrix}. \end{aligned}$$

$$\begin{aligned} f_{1:2} = P(X_2|E_{1:2}) &= \alpha \cdot O_2 \cdot T \cdot f_{1:1} \\ &= \alpha \cdot \begin{bmatrix} 0. & 0 \\ 0 & 0.5 \end{bmatrix} \cdot \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.7099 \\ 0.2901 \end{bmatrix} = \begin{bmatrix} 0.584 \\ 0.416 \end{bmatrix}. \end{aligned}$$

$$P(X_2|E_{1:2}) = \begin{bmatrix} 0.584 \\ 0.416 \end{bmatrix}.$$

Problem 8 **Making Simple Decisions (5.5 credits)**

You have to decide whether to take the theory test for a driver's license ($D \in \{d, \neg d\}$) . You can take a mock test online ($O \in \{o, \neg o\}$) before taking the real test. The result of the mock test can help you deciding whether to take the actual test. You can pass or fail the mock test ($C \in \{c, \neg c\}$) , as well as pass or fail the real test ($R \in \{r, \neg r\}$) . The following utilities are given:

$$U(o) = -40, U(\neg o) = 0, U(d, r) = 200, U(d, \neg r) = -300, U(\neg d, r) = U(\neg d, \neg r) = 0.$$

The following probabilities are given:

$$P(r|c) = 0.6, P(r|\neg c) = 0.4, P(r) = 0.56, P(c) = 0.8.$$

0 1

a) The partial ordering of the nodes for the decision network of this problem is:

Mock test (M) < Result of mock test (A) < Real test (T) < Result of real test (B)
or O < C < D < R -0.5 for each wrong node

0 1 2

b) Derive the optimal decision for D if you took the mock test, but failed. Please show your computation process in detail. No points will be given if you only present the result.

The first value set:

$$\begin{aligned} EU(t|m, \neg a) &= P(b|\neg a) \cdot U(t, m, b) + P(\neg b|\neg a) \cdot U(t, m, \neg b) \\ &= 0.3 \cdot (-200 + 500) + 0.7 \cdot (-200 + 0) - 20 \\ &= -70 \quad +1 \end{aligned}$$

$$\begin{aligned} EU(\neg t|m, \neg a) &= P(b|\neg a) \cdot U(\neg t, m, b) + P(\neg b|\neg a) \cdot U(\neg t, m, \neg b) \\ &= 0.3 * 0 + 0.7 * 0 - 20 \\ &= -20 \quad + 1 \end{aligned}$$

$$\pi^*(T|m, \neg a) = \neg t \quad \text{since} \quad EU(\neg t) > EU(t) \qquad + 0.5$$

The second value set:

$$\begin{aligned} EU(d|o, \neg c) &= P(r|\neg c) \cdot U(d, o, r) + P(\neg r|\neg c) \cdot U(d, o, \neg r) \\ &= 0.4 * (200 - 40) + 0.6 * (-300 - 40) \\ &= -140 \quad +1 \end{aligned}$$

$$\begin{aligned} EU(\neg d|o, \neg c) &= P(r|\neg c) \cdot U(\neg d, o, r) + P(\neg r|\neg c) \cdot U(\neg d, o, \neg r) \\ &= 0.4 * 0 + 0.6 * 0 - 40 \\ &= -40 \quad +1 \end{aligned}$$

$$\pi^*(D|o, \neg c) = \neg d \quad \text{since} \quad EU(\neg d) > EU(d) \qquad + 0.5$$

c) The following optimal decisions are given:

C	$\pi^*(D \mid o, C)$
c	d
$\neg c$	$\neg d$

Compute the expected utility of taking the mock test.
For readability, we present the given utilities and probabilities here again:

$$U(o) = -40, U(\neg o) = 0, U(d, r) = 200, U(d, \neg r) = -300, U(\neg d, r) = U(\neg d, \neg r) = 0.$$
$$P(r \mid c) = 0.6, P(r \mid \neg c) = 0.4, P(r) = 0.56, P(c) = 0.8.$$

The first value set:

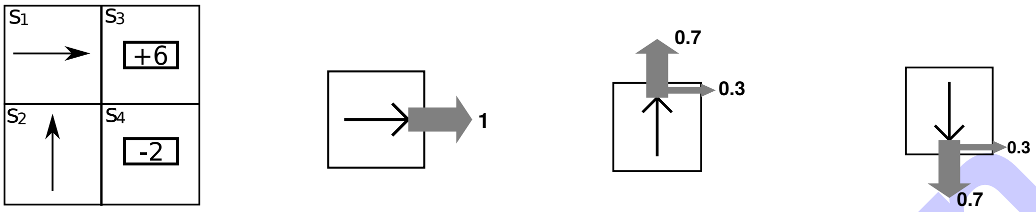
$$\begin{aligned} EU(m) = & \sum_{A \in \{a, \neg a\}} \sum_{B \in \{b, \neg b\}} P(A \mid m) \cdot P(B \mid A, m) \cdot U(\pi^*(T \mid m, A), B, m) \\ & = P(a) \cdot P(b \mid a) \cdot U(t, b, m) + P(\neg a) \cdot P(b \mid \neg a) \cdot U(\neg t, b, m) \\ & + P(a) \cdot P(\neg b \mid a) \cdot U(t, \neg b, m) + P(\neg a) \cdot P(\neg b \mid \neg a) \cdot U(\neg t, \neg b, m) \qquad +1 \\ & = 0.7 * 0.8 * (300 - 20) + 0.3 * 0.3 * (0 - 20) \\ & + 0.7 * 0.2 * (-200 - 20) + 0.3 * 0.7 * (0 - 20) \\ & = 120 \qquad +1 \end{aligned}$$

The second value set:

$$\begin{aligned} EU(o) = & \sum_{C \in \{c, \neg c\}} \sum_{R \in \{r, \neg r\}} P(C \mid o) \cdot P(R \mid C, o) \cdot U(\pi^*(D \mid o, C), R, o) \\ & = P(c) \cdot P(r \mid c) \cdot U(d, r, o) + P(\neg c) \cdot P(r \mid \neg c) \cdot U(\neg d, r, o) \\ & + P(c) \cdot P(\neg r \mid c) \cdot U(d, \neg r, o) + P(\neg c) \cdot P(\neg r \mid \neg c) \cdot U(\neg d, \neg r, o) \\ & = 0.8 * 0.6 * (200 - 40) + 0.2 * 0.4 * (0 - 40) \\ & + 0.8 * 0.4 * (-300 - 40) + 0.2 * 0.6 * (0 - 40) \\ & = -40 \qquad +1 \end{aligned}$$

Problem 9 Making Complex Decisions (6 credits)

Given is a 2x2 grid world with four states S_1, S_2, S_3, and S_4, as shown in Fig. 9.1a. The rewards of two terminal states S_3 and S_4 are $R(S_3) = 6$ and $R(S_4) = -2$, respectively. The rewards of states S_1 and S_2 are unknown $R(S_1) = R(S_2) = R$. Actions are only possible if the agent is not blocked by a wall, i.e., the possible actions at S_1 are **Right** and **Down** and the possible actions at S_2 are **Right** and **Up**. The transition probabilities of each action are shown in Fig. 9.1b, Fig. 9.1c, and Fig. 9.1d. The optimal policy is given in Fig. 9.1a. The discount factor is $\gamma = 1$.



(a) Optimal policy of a 2x2 grid world (b) Transition probability of action **Right** (c) Transition probability of action **Up** (d) Transition probability of action **Down**

Figure 9.1

0 1 2 3 4 5 6

Please derive the lower and upper bound of unknown reward R. Please round results to two digits after the decimal separator, e.g., 0.14.

The first value set of $R(S_3) = 6$ and $R(S_4) = -2$ and transition probabilities of 0.7 and 0.3

$$
\begin{aligned}
U_1 &= R(S_1) + \gamma \cdot \max_{a \in A(s)} \sum_{s'} P(s'|s,a)U(s') \\
&= R + \gamma \cdot (P(S_3|S_1) \cdot U_3 + P(S_2|S_1) \cdot U_2) \qquad +1 \\
&= R + 1 \cdot (1 \cdot 6 + 0 \cdot U_2) \\
&= R + 6 \qquad (1) \qquad +0.5
\end{aligned}
$$

$$
\begin{aligned}
U_2 &= R(S_2) + \gamma \cdot \max_{a \in A(s)} \sum_{s'} P(s'|s,a)U(s') \\
&= R + \gamma \cdot (P(S_1|S_2) \cdot U_1 + P(S_4|S_2) \cdot U_4) \qquad +1 \\
&= R + 1 \cdot (0.7 \cdot U_1 + 0.3 \cdot -2) \\
&= R + 0.7(R+6) - 0.6 \\
&= 1.7R + 3.6 \qquad (2) \qquad +0.5
\end{aligned}
$$

Optimal action at S_1 is Right :

$$
\begin{aligned}
&\implies P(S_3|S_1, Right) \cdot U_3 >= P(S_2|S_1, Down) \cdot U_2 + P(S_3|S_1, Down) \cdot U_3 \qquad +1 \\
&6 >= 0.7 \cdot U_2 + 0.3 \cdot 6 \\
&\text{Insert (2) :} \\
&4.2 >= 0.7 \cdot (1.7R + 3.6) \\
&1.19R <= 1.68 \\
&R <= 1.41 \qquad +0.5
\end{aligned}
$$

Optimal action at S_2 is Up :

$$
\begin{aligned}
&\implies P(S_1|S_2, Up) \cdot U_1 + P(S_4|S_2, Down) \cdot U_4 >= P(S_4|S_2, Right) \cdot U_4 \qquad +1 \\
&0.7 \cdot U_1 + 0.3 \cdot -2 >= 1 \cdot -2 \\
&\text{Insert (1) :} \\
&0.7 \cdot (R + 6) >= -1.4 \\
&R >= -8 \qquad +0.5
\end{aligned}
$$

The second value set of $R(S_3) = 3$ and $R(S_4) = -1$ and transition probabilities of 0.6 and 0.4

$$
\begin{aligned}
U_1 &= R(S_1) + \gamma \cdot \max_{a \in A(s)} \sum_{s'} P(s'|s,a)U(s') \\
&= R + \gamma \cdot (P(S_3|S_1) \cdot U_3 + P(S_2|S_1) \cdot U_2) \qquad +1 \\
&= R + 1 \cdot (1 \cdot 3 + 0 \cdot U_2) \\
&= R + 3 \qquad (1) \qquad +0.5
\end{aligned}
$$

$$
\begin{aligned}
U_2 &= R(S_2) + \gamma \cdot \max_{a \in A(s)} \sum_{s'} P(s'|s,a)U(s') \\
&= R + \gamma \cdot (P(S_1|S_2) \cdot U_1 + P(S_4|S_2) \cdot U_4) \qquad +1 \\
&= R + 1 \cdot (0.6 \cdot U_1 + 0.4 \cdot -1) \\
&= R + 0.6(R+3) - 0.4 \\
&= 1.6R + 1.4 \qquad (2) \qquad +0.5
\end{aligned}
$$

Optimal action at S_1 is Right :

$$
\begin{aligned}
\implies P(S_3|S_1, Right) \cdot U_3 &>= P(S_2|S_1, Down) \cdot U_2 + P(S_3|S_1, Down) \cdot U_3 \qquad +1 \\
3 &>= 0.6 \cdot U_2 + 0.4 \cdot 3 \\
\text{Insert (2) :} & \\
1.8 &>= 0.6 \cdot (1.6R + 1.4) \\
1.6R &<= 1.6 \\
R &<= 1 \qquad +0.5
\end{aligned}
$$

Optimal action at S_2 is Up :

$$
\begin{aligned}
\implies P(S_1|S_2, Up) \cdot U_1 + P(S_4|S_2, Down) \cdot U_4 &>= P(S_4|S_2, Right) \cdot U_4 \qquad +1 \\
0.6 \cdot U_1 + 0.4 \cdot -1 &>= 1 \cdot -1 \\
\text{Insert (1) :} & \\
0.6 \cdot (R+3) &>= -0.6 \\
R &>= -4 \qquad +0.5
\end{aligned}
$$

Additional space for solutions–clearly mark the (sub)problem your answers are related to and strike out invalid solutions.