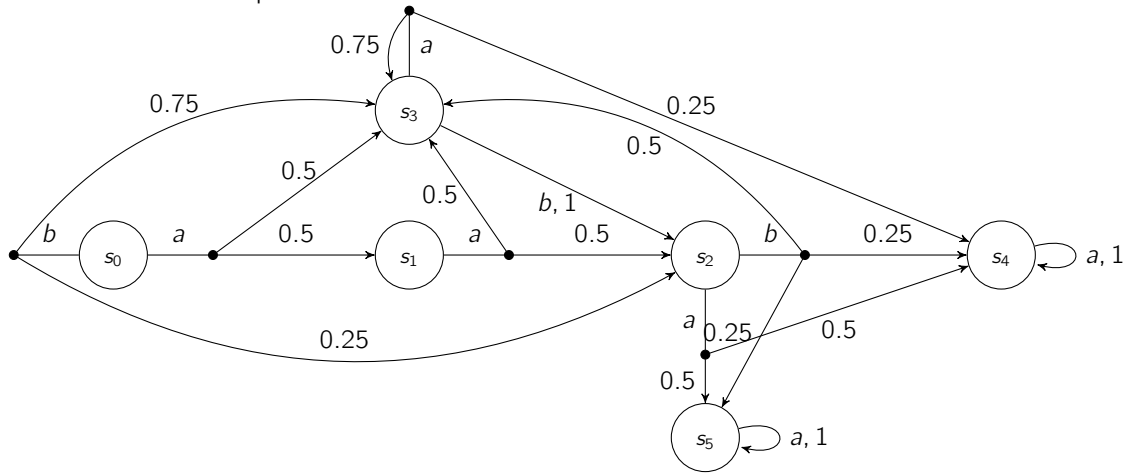


Exercise 1 (Reachability Probabilities in MDPs):

(5 points)

Consider the MDP depicted below.



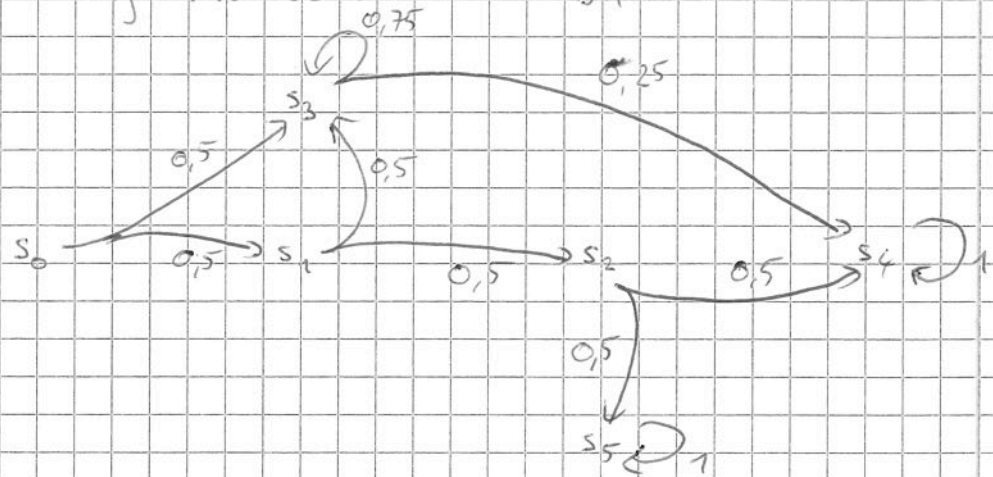
Consider the probabilities for the following property: $\Pr_{\max}^{\mathcal{M}}(s \models \Diamond s_4)$ for all $s \in S$.

- Execute policy iteration, start with taking a in all states. If you need an ordering on the states, use the ordering of the state numbers.
- Give the LP formulation to characterize the property above.
- Execute the preprocessing to eliminate $S^{=0}, S^{=1}$ as discussed in the lecture. Then, execute two iterations of value iteration.

Solution: _____

1a) Choosing α in all states yields the following induced DTMC $\mathcal{D}_1$

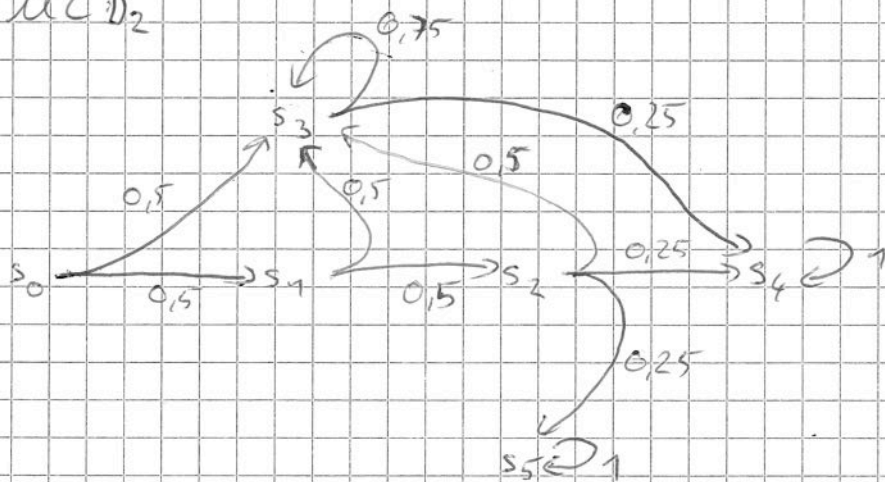
(1)



Solving this DTMC yields

$$Pr^{D_1}(s_i = \bigcirc s_4) = \begin{cases} 1 & \text{if } i=3 \text{ or } i=4 \\ 0 & \text{if } i=5 \\ 0.5 & \text{if } i=2 \\ 0.75 & \text{if } i=1 \\ 0.875 & \text{if } i=0 \end{cases}$$

Hence, we should improve the policy at state s_2 by choosing b instead of a . All other choices are locally optimal. We get the following induced DTMC $\mathcal{D}_2$



We obtain the following probabilities:

(2)

$$Pr^{D_4}(s_i \neq \Diamond s_4) = \begin{cases} 1 & \text{if } i = 3 \text{ or } i = 4 \\ 0 & \text{if } i = 5 \\ 0.75 & \text{if } i = 2 \\ 0.875 & \text{if } i = 1 \\ 0.9375 & \text{if } i = 0 \end{cases}$$

Since no decision can be locally improved, policy iteration terminates and $Pr^{\max}(s_i \neq \Diamond s_4)$ coincides with $Pr^{D_2}(s_i \neq \Diamond s_4)$.

b) In fact, for computing maximal probabilities no preprocessing is necessary for the LP and we get: (3)

$$\text{minimize } \sum_{s \in S} x_s \quad \text{subject to}$$

$$x_{s_4} = 1 \quad \text{if } s = s_4$$

$$x_{s_5} \geq 1 \cdot x_{s_5}$$

$$x_{s_2} \geq 0,5 \cdot x_{s_4} + 0,5 \cdot x_{s_5}$$

$$x_{s_2} \geq 0,5 \cdot x_{s_3} + 0,25 \cdot x_{s_4} + 0,25 \cdot x_{s_5}$$

$$x_{s_6} \geq 0,75 \cdot x_{s_3} + 0,25 \cdot x_{s_4}$$

$$x_{s_3} \geq 1 \cdot x_{s_2}$$

$$x_{s_1} \geq 0,5 \cdot x_{s_2} + 0,5 \cdot x_{s_5}$$

$$x_{s_0} \geq 0,5 \cdot x_{s_1} + 0,5 \cdot x_{s_3}$$

$$x_{s_0} \geq 0,25 \cdot x_{s_2} + 0,75 \cdot x_{s_3}$$

c) For maximal probabilities, we have

$$S^0 = \{s \mid Pr^{\max}(s \models \Diamond s_4) = 0\}$$

$$= \{s \mid s \models \neg \exists \Diamond s_4\}$$

$$= \{s_5\}$$

In order to compute

$$S^1 = \{s \mid Pr^{\max}(s \models \Diamond s_4) = 1\}$$

we start from S^0 and iteratively remove predecessor actions and states (if they have no outgoing action anymore).

(i) remove $\{a, b\}$ from $Act(s_2)$, because

$$Post(s_2, a) \cap S^0 \neq \emptyset \text{ and}$$

$$Post(s_2, b) \cap S^0 \neq \emptyset.$$

(ii) remove s_2 since $Act(s_2) = \emptyset$

(iii) remove a from $Act(s_1)$, because

$$Post(s_1, a) \cap \{s_2, s_5\} \neq \emptyset$$

(iv) remove s_1 since $Act(s_1) = \emptyset$

(v) remove b from $Act(s_3)$, because

$$Post(s_3, b) \cap \{s_2, s_5\} \neq \emptyset$$

(vi) remove b from $Act(s_0)$, because

$$Post(s_0, b) \cap \{s_2, s_5\} \neq \emptyset$$

(vii) remove a from $Act(s_0)$, because

$$Post(s_0, a) \cap \{s_1, s_2, s_5\} \neq \emptyset$$

(viii) remove s_0 since $Act(s_0) = \emptyset$

Since the remaining MDP has no states that can not reach s_4 , we are done and obtain

$$S^1 = \{s_3, s_4\}$$

Executing two steps of value iteration proceeds as follows:

⑤

Initial probability vector

$$(0 \ 0 \ 0 \ \underbrace{1 \ 1 \ 0})$$

fixed by pre processing

First step:

$$(\alpha \ \beta \ \gamma \ 1 \ 1 \ 0)$$

where

$$\alpha = \max(0,5 \cdot 0 + 0,5 \cdot 1, 0,25 \cdot 0 + 0,75 \cdot 1) = 0,75$$

$$\beta = \max(0,5 \cdot 0 + 0,5 \cdot 1) = 0,5$$

$$\gamma = \max(0,5 \cdot 1 + 0,5 \cdot 0, 0,5 \cdot 1 + 0,25 \cdot 1 + 0,25 \cdot 0) = 0,75$$

$$\Rightarrow (0,75 \ 0,5 \ 0,75 \ 1 \ 1 \ 0)$$

Second step:

$$(\alpha \ \beta \ \gamma \ 1 \ 1 \ 0)$$

where

$$\alpha = \max(0,5 \cdot 1 + 0,5 \cdot 0,5, 0,25 \cdot 0,75 + 0,75 \cdot 1) = 0,9375$$

$$\beta = \max(0,5 \cdot 0,75 + 0,5 \cdot 1) = 0,875$$

$$\gamma = \max(0,5 \cdot 1 + 0,5 \cdot 0, 0,5 \cdot 1 + 0,25 \cdot 1 + 0,25 \cdot 0) = 0,75$$

$$\Rightarrow (0,9375 \ 0,875 \ 0,75 \ 1 \ 1 \ 0)$$

No changes are made in the next step and value iteration converged.

Exercise 2 (Minimal Counterexamples):

(5 points)

For an MDP $\mathcal{M} = (S, \text{Act}, \mathbf{P}, s_{\text{init}}, \text{AP}, \mathbf{L})$, a *subsystem* is an MDP $\mathcal{M}' = (S', \text{Act}', \mathbf{P}', s'_{\text{init}}, \text{AP}', \mathbf{L}')$ with $S' \subset S$. $\mathcal{M}'$ shall have the exact same behavior as $\mathcal{M}$ for those states, that are included in S' .

- a) Give a formal definition for a subsystem of an MDP. What changes need to be made to the original Definition of MDPs?
- b) A *mixed integer linear program* is a linear program where certain variables are allowed to be integer. The formal definition reads as follows: Let $A \in \mathbb{Q}^{m \times n}$, $B \in \mathbb{Q}^{m \times k}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}^n$, and $d \in \mathbb{Q}^k$. A *mixed integer linear program* (MILP) consists of computing $\min c^T x + d^T y$ such that $Ax + By \leq b$ and $x \in \mathbb{R}^n$, $y \in \mathbb{Z}^k$. Consider an MDP $\mathcal{M} = (S, \text{Act}, \mathbf{P}, s_{\text{init}}, \text{AP}, \mathbf{L})$ with a single initial state and a reachability property $\varphi = \mathbb{P}_{\leq \lambda}(\Diamond \text{target})$ that is violated for s_{init} .

Give an MILP formulation that computes a subsystem $\mathcal{M}'$ of an MDP $\mathcal{M}$ which is minimal in terms of the number of states such that s_{init} is included in S' and φ is also violated for s_{init} inside $\mathcal{M}'$. (*Hint: Use integer variables to count the states of the subsystem.*)

Solution: _____

Series 6, Solution

1

~~11/12~~

EA Given: MDP $M = (S, Act, P, l_{init}, AP, L)$

Subsystem: MDP $M' = (S', s'_{init}, Act', P', L')$ is subsystem of M if $S' \subseteq S$, $s'_{init} = s_{init}$, $L'(s) = L(s)$ for all $s \in S'$, $Act' \subseteq Act$, $P'(s, \alpha, s') \neq 0 \Rightarrow P'(s, \alpha, s') = P(s, \alpha, s')$ for all $s, s' \in S'$ and $\alpha \in Act'$.

minimization:

introduce: $x_s \in [0, 1] \subseteq \mathbb{R}$ for all $s \in S$
 $p_s \in [0, 1] \subseteq \mathbb{R}$ "
 $\sigma_{s, \alpha} \in [0, 1] \subseteq \mathbb{R}$ for all $s \in S, \alpha \in Act$

minimize

$$\sum_{s \in S} x_s$$

2

s.t. $p_{s_T} > \lambda$

prob of initial state shall at least exceed λ

$\forall s \in T: p_s = x_s$

if target included, prob will be 1

$\forall s \in SIT: \sum_{\alpha \in Act} \sigma_{s,\alpha} = x_s$

if s included, exactly 1 action is chosen

$\forall s \in SIT \forall \alpha \in Act: p_s \leq (1 - \sigma_{s,\alpha}) + \sum_{s' \in succ_n(s,\alpha)} P(s,\alpha,s') \cdot p_{s'}$

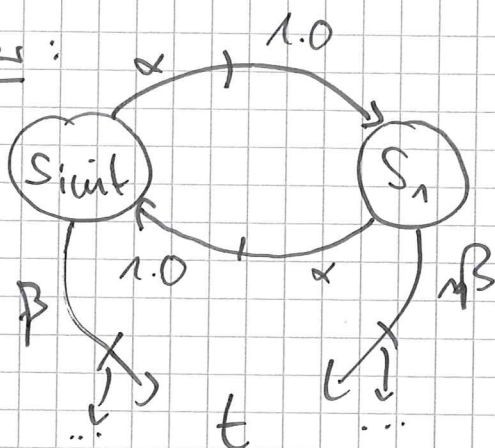
if α at s ,
 $= 0$, otherwise
 $= 1$, trivially
 fulfills
 constraint

reachability prob
 of s

2. Question: Is this correct?

3

Consider:

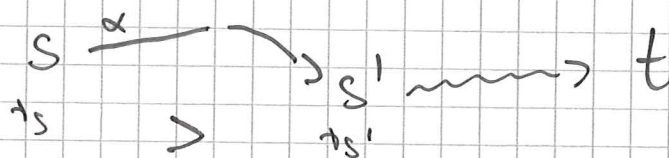


- selecting α in both states prevents reaching the target states $\rightarrow$ problematic states

$$P_{S_{init}} = 1 \cdot P_{S_1} \quad P_{S_1} = 1 \cdot P_{S_{init}}$$

for all $\alpha \in [0, 1]$: $P_{S_{init}} = P_{S_1} = \alpha$ is a solution,
although t is not reachable.

Solution: *Variable $\tau_s \in \mathbb{R}$ to problematic states.



not possible, strictly increasing!

$\rightarrow$ ensures existence of loop-free path from problematic state to t !

Encoding:

4

$$\text{Let } \text{Act}_M^{\text{Probl}} = \{(s, \alpha) \in S \times \text{Act} \mid \text{succ}_M(s, \alpha) \subseteq S_M^{\text{Probl}}\}$$

$$s \xrightarrow{\alpha} \begin{matrix} s_1 \\ s_1 \end{matrix} \} \text{problematic!}$$

(only problematic states reachable via s and α !)

$$\forall s \in S^{\text{Probl}}: \tau_s \in \mathbb{R} \text{ (partial order on problematic states)}$$

$$\forall s, s' \in S^{\text{Probl}}: t_{ss'} \in \mathbb{R}^B: \text{transition in the MCS between states } s, s'$$

$$\forall (s, \alpha) \in \text{Act}^{\text{Probl}} \forall s' \in \text{succ}(s, \alpha): t_{ss'} \leq x_{s'}$$

if s' in MCS, transition might exist

$$\forall (s, \alpha) \in \text{Act}^{\text{Probl}} \forall s' \in \text{succ}(s, \alpha): \tau_s < \tau_{s'} + (1 - t_{ss'})$$

partial order

if transition is chosen

$$\forall (s, \alpha) \in \text{Act}^{\text{Probl}}: (1 - \sigma_{s, \alpha}) + \sum_{s' \in \text{succ}(s, \alpha)} t_{ss'} \geq x_s$$

if scheduler assigns α at s

1) if s chosen, transition is needed

$$s \rightarrow s' \rightarrow s'' \rightarrow \dots \rightarrow t$$

$$\tau_s < \tau_{s'} < \tau_{s''} < \dots$$

not τ_t !

"escapes at target states!"