רן מזר - 315489765

חגי קרזנסקי - 316119668,

(1)

(i) $V_\pi(s) = E\left[R_{t+1} + \gamma V_\pi(S_{t+1}) \mid S_t = s\right] = \sum_{s' \in S} Tr(s,a,s')\left(R(s,\pi(s),s') + \gamma \underbrace{V_\pi(s')}\right)$

(ii) $q_\pi(s,a) = \sum_{s' \in S} Tr(s,a,s')\left(R(s,\pi(s),s') + \gamma q_\pi(s',\pi(s'))\right)$

optimally

(iii) $V_*(s) = \max_\pi V_\pi(s) = \max_\pi \sum_{s' \in S} Tr(s,a,s')\left(R(s,\pi(s),s') + \gamma V_\pi(s')\right)$

(iv) $q_*(s,a) = \underbrace{\max_\pi}\, q_\pi(s) = \max_\pi \sum_{s' \in S} \underbrace{Tr(s,a,s')}\left(R(s,\pi(s),s') + \gamma q_\pi(s',\pi(s'))\right)$

(2) ראשית נשים לב כי כמות המדיניות (π) היא סופית $|A|^{|S|}$.

לפי האלגו' (modified π iteration) ראינו בהרצאה כי

- It therefore improves the value function, $v_{\pi'}(s) \geq v_\pi(s)$

$$\begin{aligned} v_\pi(s) &\leq q_\pi(s,\pi'(s)) = \mathbb{E}_{\pi'}\left[R_{t+1} + \gamma v_\pi(S_{t+1}) \mid S_t = s\right] \\ &\leq \mathbb{E}_{\pi'}\left[R_{t+1} + \gamma q_\pi(S_{t+1},\pi'(S_{t+1})) \mid S_t = s\right] \\ &\leq \mathbb{E}_{\pi'}\left[R_{t+1} + \gamma R_{t+2} + \gamma^2 q_\pi(S_{t+2},\pi'(S_{t+2})) \mid S_t = s\right] \\ &\leq \mathbb{E}_{\pi'}\left[R_{t+1} + \gamma R_{t+2} + \ldots \mid S_t = s\right] = v_{\pi'}(s) \end{aligned}$$

ולכן הכי שנשפר את המדיניות היא מונוטונית עולה וסופית <= חסומה

<= ולכן ידוע שהיא מתכנסת. ובפרט מקבלת מקסימום (כי היא סופית)

הסדרה שנקבל מריצת האלגו' מתכנסת <= לכל $s \in S$ -> $V_i(s)$ היא סדרת קושי.

<= קיים i כך ש: לכל $\delta > 0$: $|V_i^\pi - V_{i+1}^\pi| < \delta$

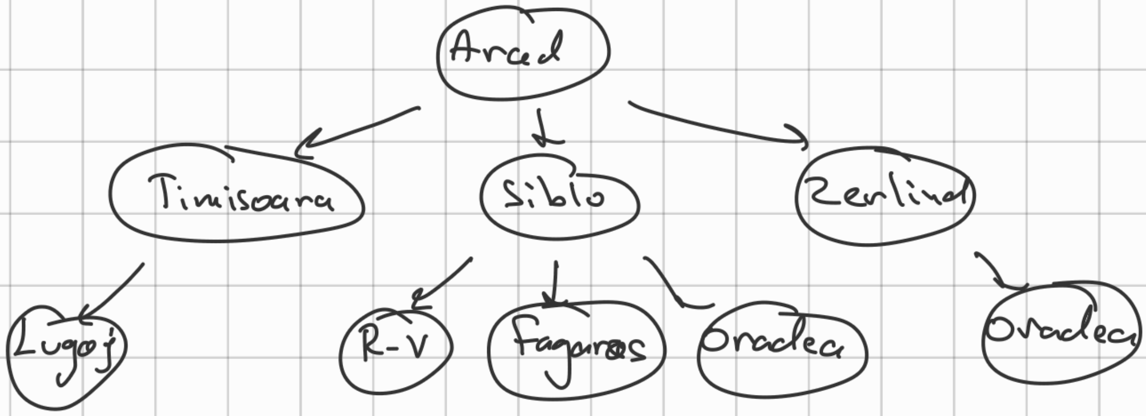
כמו כן ידוע שיש f כך שלכל $\varepsilon > 0$, $f(\varepsilon) = \delta$

<= $|V_i^\pi - V^{\pi_*}|_\infty < \varepsilon$

3) סעיף 3.

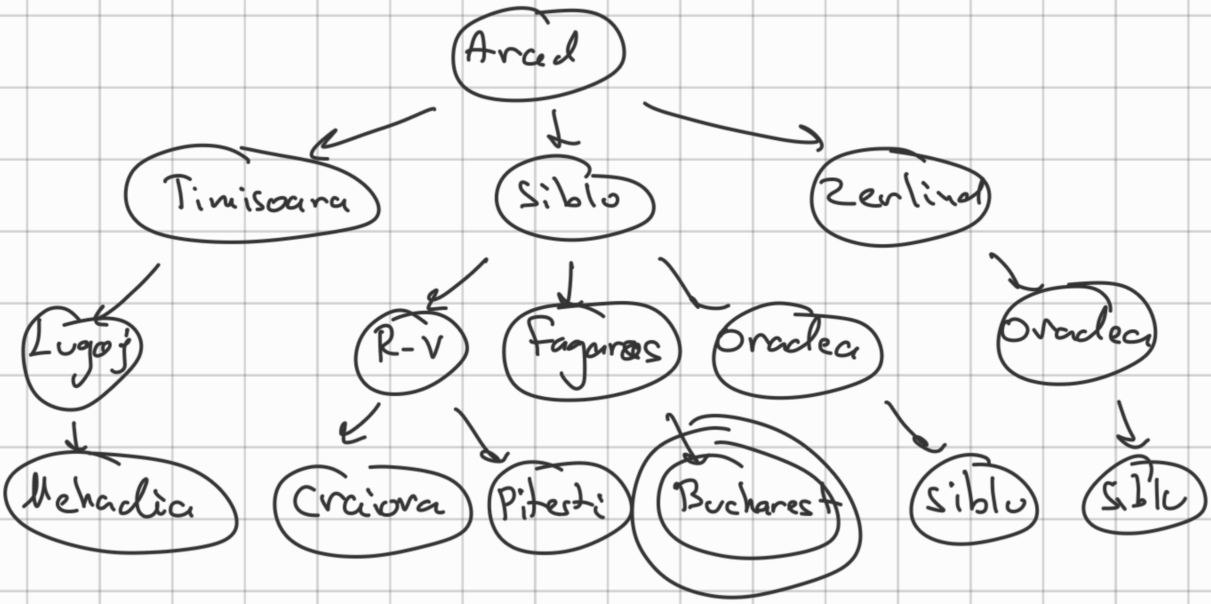
א) עומק 2:

(3 איטרציות)



עומק 3:

(4 איטרציות)



ח) א)

At (x, y)

holding (x)

Empty (robot) = ∀x ∈ toys not holding (x)

ב)

move (x, y)

x, y → locations (room)

pre → At (robot, x)

post → At (robot, y)

pickUp (x)

x → color   y → room

pre → At (robot, y) ∧ At (x, y) ∧ Empty (robot)

post → holding (x), not Empty (robot), At(robot, y), At(x, y).

put Down (x)

$x \to$ color

pre $\to$ $At(robot, y) \wedge At(x, y) \wedge$
holding(x)

post $\to$ Empty(robot), At(robot, y), At(x, y)

Deliver(x, y)

$x \to$ color, $y \to$ child

pre $\to$ z - children room, At(robot, z), At(y, z),
holding(x),

post $\to$ z - children room, At(robot, z), At(y, z),
Empty(robot), At(x, y).

Empty(robot) ②

At(robot, room1)

At(child, children room)

$\forall x \in$ colors : At(x, room 1).

x_1 = blue, x_2 = red ③

$At(x_1, child) \wedge At(x_2, child)$