Logic, Learning, and Decision

Course code: SSY165

Examination 2023-10-23

Time: 8:30-12:30, Location: Johanneberg

Teacher: Bengt Lennartson, phone 0730-79 42 26

The examination includes 25 points, where grade three requires 10 points, grade four 15 points and grade five 20 points.

The result of this examination is announced and inspection of the grading is done on November 9 and 10, 12:30-13:00 at the division.

Allowed aids at the examination:

- Standard mathematical tables such as Beta, see also formulas in the end of this examination.
- Pocket calculator.

Good luck!

Department of Electrical Engineering
Division of Systems and Control
Chalmers University of Technology

1

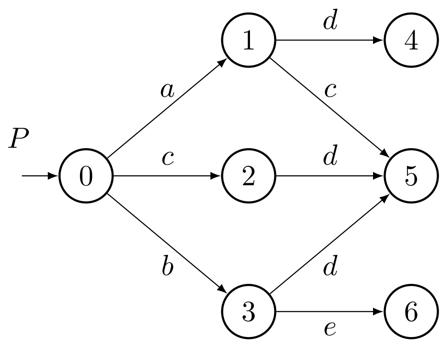
Prove the following set equivalence by equivalence relations based on predicate expressions:

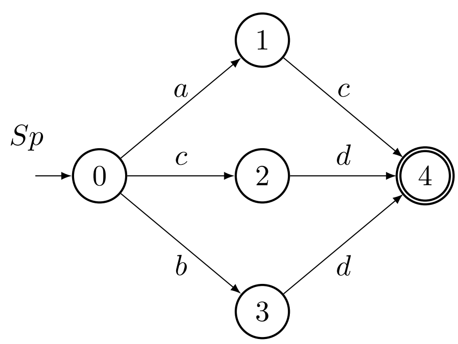
$$A \cap B \subseteq B \cap C \Leftrightarrow A \cap C \cap B = A \cap B$$

(3 p)

2

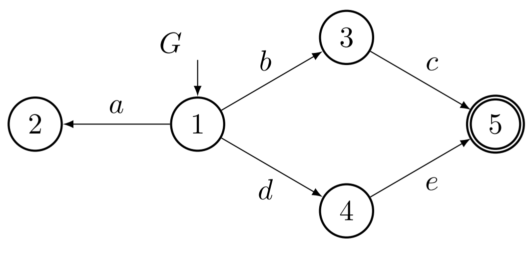
For the plant P and specification Sp below, generate a controllable and nonblocking supervisor when the events d and e are uncontrollable, while the events a, b and c are controllable. Show the resulting automaton after each Backward_Reachability (Coreachability) computation. (4 p)





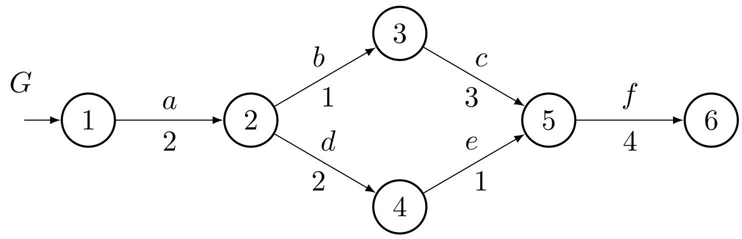
3

A nonblocking state satisfies the CTL expression $\forall\Box\exists\Diamond p$. Show by μ-calculus that the following automaton is blocking. (4 p)



4

Consider the automaton G, including immediate rewards on each transition.



a) Iterate the Q-learning algorithm

$$\widehat{Q}_{k+1}(x,a) = (1-\alpha_k)\widehat{Q}_k(x,a) + \alpha_k\big(r' + \gamma \max_{b\in\Sigma(x')} \widehat{Q}_k(x',b)\big)$$

until convergence for $\alpha_k = 1$ and $\gamma = 1$. The action with the largest estimated $\widehat{Q}$-value is chosen (greedy action) in state 2, except when $\widehat{Q} = 0$, which has higher priority. This strategy is chosen to improve the initial exploration. (3 p)

b) Based on the resulting $\widehat{Q}$-function in a), determine the optimal sequence of actions that maximizes the total reward. Confirm this result by computing the optimal total reward and optimal sequence by dynamic programming. (2 p)

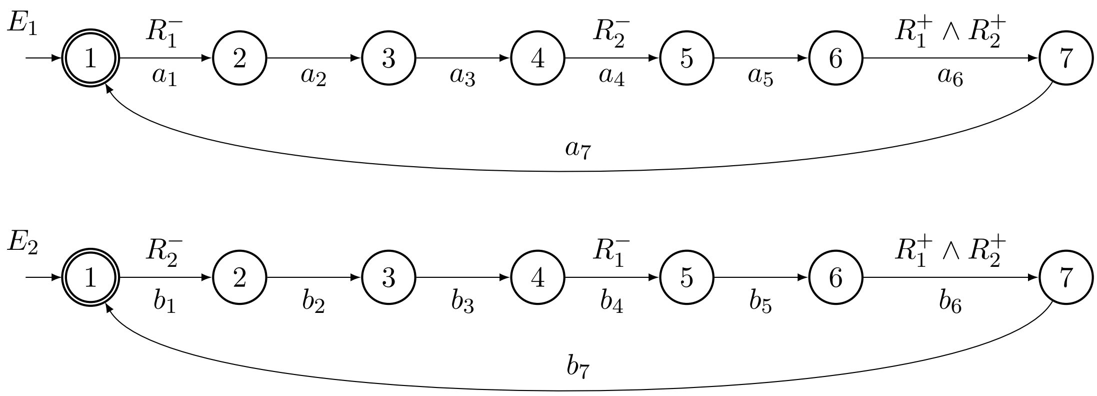
5

In an initial state, two alternative actions can be chosen. Taking action a implies that the desired next state is always reached with a reward 1, followed by an immediate transition back to the initial state without any additional reward. Taking action b gives a higher reward r, but with a risk to reach another next state with negative reward -10. The successful outcome has a transition probability 0.8, while the state with negative reward is reached with transition probability 0.2. In both cases an immediate transition back to the initial state occurs without any additional reward.

Which is the lowest successful reward r that is required to make it profitable on average to select the uncertain action with the risk to obtain a negative reward? (4 p)

6

Consider the following two extended finite automata with the shared variables R_i, $i = 1, 2$. The notion $R_i^{\pm}$ at a transition means that the updated value of R_i after such a transition is $R_i' = R_i \pm 1$. The transition is however only admissible if the next value is 0 or 1. The initial value is $R_i = 1$.



To reduce the number of states, local events that are not involved in any other subsystems can be replaced by the hidden event τ. Any τ transition can then be removed when no alternative transitions are involved in the source state of such τ transitions, and the source and target states have the same state label. Removing a τ transition means that the source and target states are merged into one state.

a) Reformulate the synchronous composition $E_1 \,\|\, E_2$ as a synchronization of four ordinary automata

$$G_1 \,\|\, G_2 \,\|\, G_3 \,\|\, G_4,$$

where the automata G_3 and G_4 model the two variables R_1 and R_2 and their interaction with the two sequences in E_1 and E_2. The two sequences, now interacting with G_3 and G_4 by shared events, are represented by the automata G_1 and G_2. (2 p)

b) Apply the state reduction principle mentioned above and then compute

$$(G_1^{\mathcal{A}} \,\|\, G_2^{\mathcal{A}} \,\|\, G_3 \,\|\, G_4)^{\mathcal{A}}$$

where the reduction (abstraction operator) $\mathcal{A}$ includes the replacement of local events with τ, followed by the state reduction. Only three states will remain if in the final step the additional principle is applied that two states with the same future behavior can be joined into one state block. Also note that all events are local τ events when all four automata have been synchronized. (3 p)

Solution Exam Logic, Learning & Decision
23/023

1. Prove that $A \cap B \subseteq B \cap C \Leftrightarrow A \cap C \cap B = A \cap B$

Since $A \cap C \cap B = A \cap B \cap B \cap C$ introduce the notations $D \overset{def}{=} A \cap B$ and $E \overset{def}{=} B \cap C$

Now it is enough to prove that

$$D \subseteq E \Leftrightarrow D \cap E = D$$

$$D \cap E = D \Leftrightarrow D \cap E \subseteq D \wedge D \subseteq D \cap C \Leftrightarrow$$

$$\forall x : (x \in D \wedge x \in E \rightarrow x \in D) \wedge (x \in D \rightarrow (x \in D \wedge x \in E)) \Leftrightarrow$$

$$\forall x : (x \notin D \vee x \notin E \vee x \in D) \wedge (x \notin D \vee (x \in D \wedge x \in E)) \Leftrightarrow$$

$$\forall x : (\underbrace{x \notin D \vee x \in D}_{\pi} \vee x \notin D) \wedge (\underbrace{x \notin D \vee x \in D}_{\pi}) \wedge (x \notin D \vee x \in E) \Leftrightarrow$$

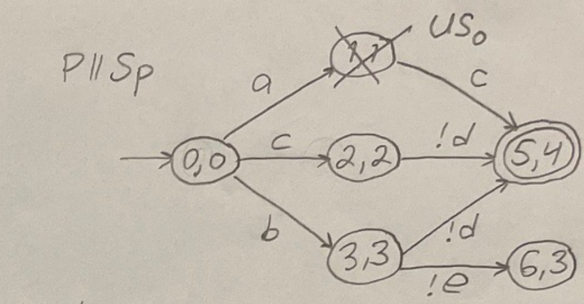
$$\forall x : \underbrace{(\pi \vee x \notin D)}_{\pi} \wedge \pi \wedge \pi \wedge (x \in D \rightarrow x \in E) \Leftrightarrow$$

$$\forall x : \pi \wedge (x \in D \rightarrow x \in E) \Leftrightarrow$$

$$\forall x : x \in D \rightarrow x \in E \Leftrightarrow$$

$$D \subseteq E$$

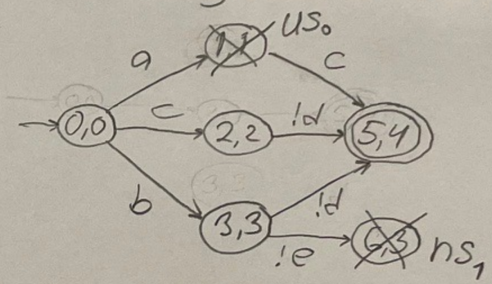
2.



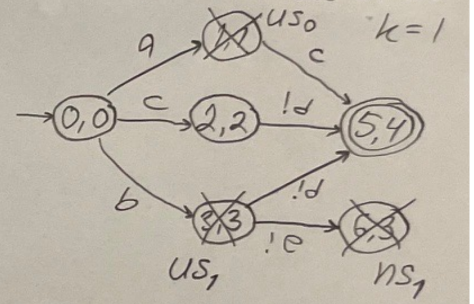
US_k = uncontrollable state in iteration k

bS_k = blocking state in iteration k

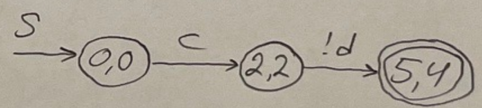
Nonblocking states k=1



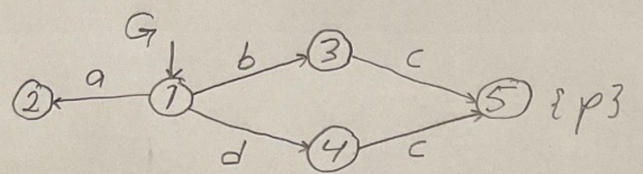
Uncontrollable states k=1



Fixpunkt in iteration 2 gives the supervisor



3.



Show by μ-calculus that

$G \not\models \forall\Box\exists\Diamond p \Leftrightarrow \{1\} \notin [\![\forall\Box\exists\Diamond p]\!]$

$[\![\exists\Diamond p]\!] = [\![\mu y. p \vee \exists\circ y]\!] = \mu Y. \Psi(Y)$

$\Psi(Y) = [\![p]\!] \cup Pre^{\exists}(Y) \qquad Y_{k+1} = \Psi(Y_k) \qquad Y_0 = \emptyset$

$Y_1 = \{5\} \cup Pre^{\exists}(\emptyset) = \{5\}$

$Y_2 = \{5\} \cup Pre^{\exists}(\{5\}) = \{5\} \cup \{3,4\} = \{3,4,5\}$

$Y_3 = \{5\} \cup Pre^{\exists}(\{3,4,5\}) = \{5\} \cup \{1,3,4\} = \{1,3,4,5\}$

$Y_4 = Y_3 = Y^{\omega} = \{1,3,4,5\} \qquad [\![\exists\Diamond p]\!] = Y^{\omega} = \{1,3,4,5\}$

$[\![\forall\Box\exists\Diamond p]\!] = [\![\forall\Box w]\!]$ where $[\![w]\!] = Y^{\omega}$

$[\![\forall\Box z]\!] = [\![\nu z. w \wedge \exists\circ z]\!] = \nu Z. \Psi(Z)$

$\Psi(Z) = [\![w]\!] \cap Pre^{\forall}(Z) \qquad Z_{k+1} = \Psi(Z_k), \quad Z_0 = X$

$Z_1 = [\![w]\!] \cap Pre^{\forall}(X) = \{1,3,4,5\} \cap X = \{1,3,4,5\}$

$Z_2 = [\![1,3,4,5]\!] \cap Pre^{\forall}(\{1,3,4,5\}) = \{3,4,5\}$

1 is excluded since its target state 2 is not included in Z_1.

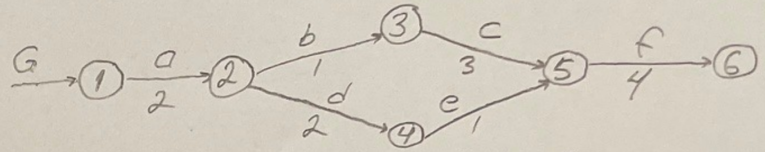
$Z_3 = [\![1,3,4,5]\!] \cap Pre^{\forall}(\{3,4,5\}) = \{3,4,5\} = Z_2 = Z^{\omega}$

5 is included since $\forall a \in \Sigma(x): p(x,a)$ is true for any $p(x,a)$ when $\Sigma(x) = \emptyset$.

$\because \{1\} \notin [\![\forall\Box\exists\Diamond p]\!] = \{3,4,5\} \Rightarrow G \not\models \forall\Box\exists\Diamond p$

$\Leftrightarrow$ G is blocking. From state 2 state 5 cannot be reached

4.



a) $\hat{Q}_{k+1}(x,a) = r' + \max_{b \in \Sigma(x')} \hat{Q}_k(x',b) \qquad \hat{Q}_0(x,a) = 0$

k	x	a	x'	r'	$\hat{Q}_k(x',b)$	$\hat{Q}_{k+1}(x,a)$
0	1	a	2	2	$\hat{Q}(2,b) = \hat{Q}(2,d) = 0$	$\hat{Q}(1,a) = 2$
1	2	b	3	1	$\hat{Q}(3,c) = 0$	$\underline{\hat{Q}(2,b) = 1}$
2	3	c	5	3	$\hat{Q}(5,f) = 0$	$\underline{\hat{Q}(3,c) = 3}$
3	5	f	6	4	$\hat{Q}(6,-) = 0$	$\hat{Q}(5,f) = 4$
4	1	a	2	2	$\hat{Q}(2,b) = 1,\ \hat{Q}(2,d) = 0$	$\hat{Q}(1,a) = 2+1 = 3$
5	2	d	4	2	$\hat{Q}(4,e) = 0$	$\hat{Q}(2,d) = 2$
6	4	e	5	1	$\hat{Q}(5,f) = 4$	$\hat{Q}(4,e) = 1+4 = 5$
7	5	f	6	4	$\hat{Q}(6,-) = 0$	$\underline{\hat{Q}(5,f) = 4}$
8	1	a	2	2	$\hat{Q}(2,b) = 1,\ \hat{Q}(2,d) = 2$	$\hat{Q}(1,a) = 2+2 = 4$
9	2	d	4	2	$\hat{Q}(4,e) = 5$	$\hat{Q}(2,d) = 2+5 = 7$
10	4	e	5	1	$\hat{Q}(5,f) = 4$	$\underline{\hat{Q}(4,e) = 1+4 = 5}$
11	5	f	6	4	$\hat{Q}(6,-) = 0$	$\hat{Q}(5,f) = 4$
12	1	a	2	2	$\hat{Q}(2,b) = 1,\ \hat{Q}(2,d) = 7$	$\hat{Q}(1,a) = 2+7 = 9$
13	2	d	4	2	$\hat{Q}(4,e) = 5$	$\underline{\hat{Q}(2,d) = 2+5 = 7}$
14	4	e	5	1	$\hat{Q}(5,f) = 4$	$\hat{Q}(4,e) = 5$
15	5	f	6	4	$\hat{Q}(6,-) = 0$	$\hat{Q}(5,f) = 4$
16	1	a	2	2	$\hat{Q}(2,b) = 1,\ \hat{Q}(2,d) = 7$	$\underline{\hat{Q}(1,a) = 2+7 = 9}$

Underlined $\hat{Q}_{k+1}(x,a)$ value has converged to a constant value. It does not change anymore.

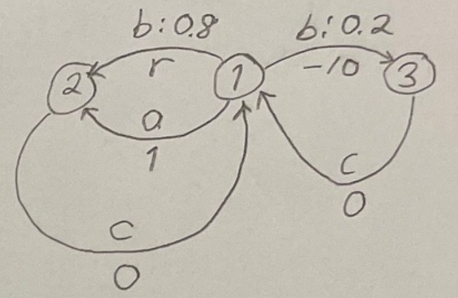
b) Since the convergent value of $\hat{Q}(2,d)=7 > \hat{Q}(2,b)=1$, the greedy strategy means that action d is taken in state 2, which generates the sequence $1 \xrightarrow{a} 2 \xrightarrow{d} 4 \xrightarrow{e} 5 \xrightarrow{f} 6$

Dynamic programming gives on the other hand $J^*(5)=4$; $J^*(3)=3+4=7$, $J^*(4)=1+4=5$, $J^*(2)=\max\{1+J^*(3), 2+J^*(4)\} = \max\{1+7, 2+5\} = \max\{8,7\} = 8$. This result is achieved by taking the action b in state 2. $J^*(1)=2+8=10$ i.e. the optimal sequence is

$1 \xrightarrow{a} 2 \xrightarrow{b} 3 \xrightarrow{c} 5 \xrightarrow{f} 6$

The incorrect greedy $\hat{Q}$-solution is achieved because not enough of exploration (no ε actions) is involved in the solution.

5.



action a

$$[p_1\ p_2\ p_3] = [p_1\ p_2\ p_3]\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} =$$

$$= [p_2+p_3\ \ p_1\ \ 0] \Rightarrow \begin{cases} p_1 = p_2+p_3 \\ p_2 = p_1 \\ p_3 = 0 \end{cases}$$

$p_1 = p_2$, $p_3 = 0$, $p_1+p_2+p_3 = 1 \Rightarrow 2p_1 = 1$

$p_1 = p_2 = 0.5$ $\quad J_a = 1\, p_2 = 0.5$

action b

$$[p_1\ p_2\ p_3] = [p_1\ p_2\ p_3]\begin{bmatrix} 0 & 0.8 & 0.2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} =$$

$$= [p_2+p_3\ \ 0.8p_1\ \ 0.2p_1] \Rightarrow \begin{cases} p_1 = p_2+p_3 \\ p_2 = 0.8p_1 \\ p_3 = 0.2p_1 \end{cases}$$

$p_1+p_2+p_3 = p_1+0.8p_1+0.2p_1 = 1$

$p_1 = 0.5$, $p_2 = 0.8p_1 = 0.4$ $\quad p_3 = 0.2p_1 = 0.1$

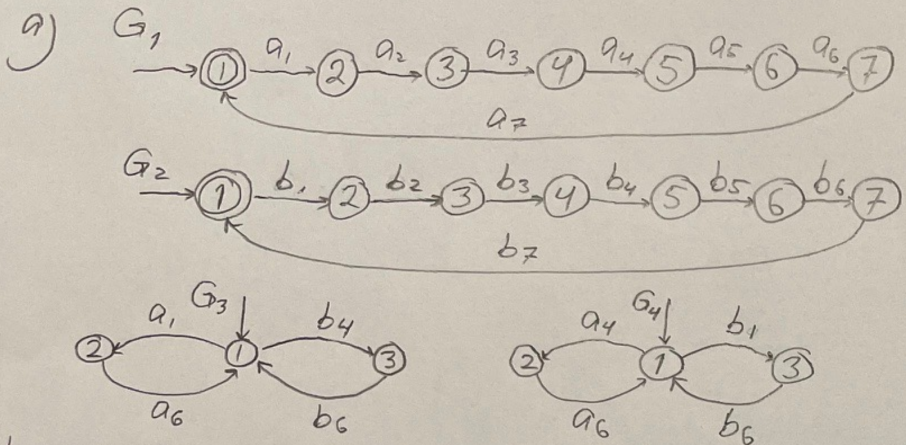
$J_b = r\,p_2 - 10\,p_3 = 0.4r - 10\cdot 0.1 = 0.4r - 1$

$J_b > J_a \Rightarrow 0.4r - 1 > 0.5 \Leftrightarrow r > \frac{1.5}{0.4} = 3.75$

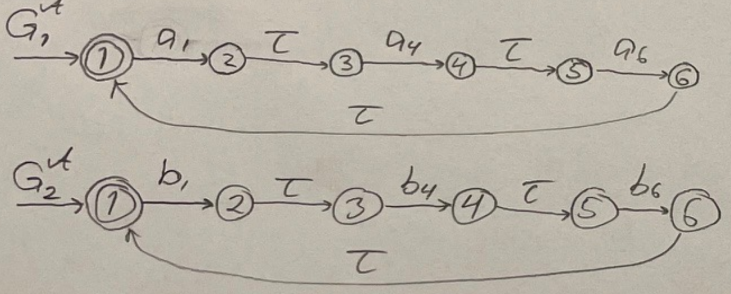
∴ $r > 3.75$ means that on average it is more profitable to select the uncertain action.

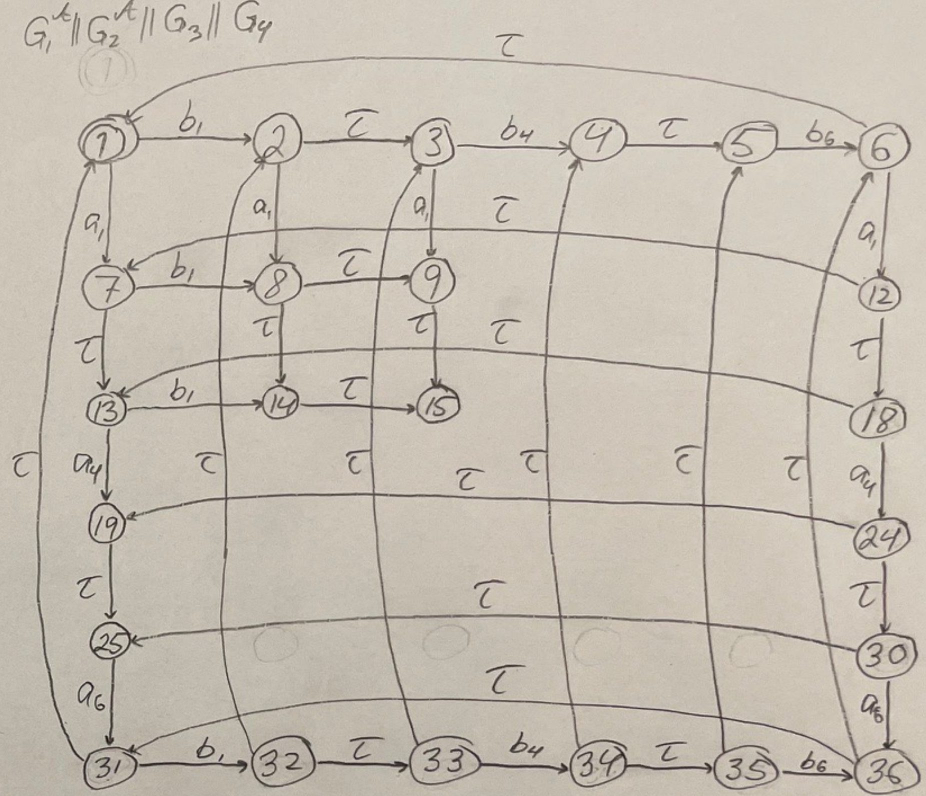
6.

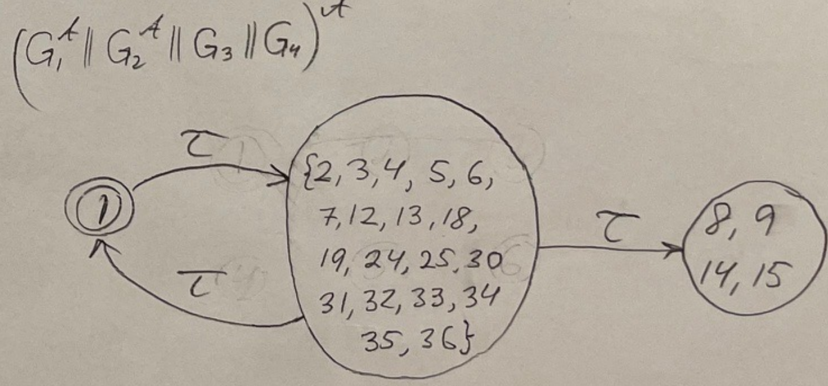
a)


G1
a1 a2 a3 a4 a5 a6
a7
G2
b1 b2 b3 b4 b5 b6
b7
G3
a1 b4
a6 b6
G4
a4 b1
a6 b6

b)


G1^A
a1 τ a4 τ a6
τ
G2^A
b1 τ b4 τ b6
τ


G1^A || G2^A || G3 || G4
b1 τ b4 τ b6
τ
a1 b1 τ
τ b1 τ
a4
a6
31 b1 32 τ 33 b4 34 τ 35 b6 36


(G1^A || G2^A || G3 || G4)^A
τ
{2, 3, 4, 5, 6, 7, 12, 13, 18, 19, 24, 25, 30 31, 32, 33, 34 35, 36}
τ
8, 9 14, 15
τ