The Cayley-Hamilton Theorem For Finite Automata

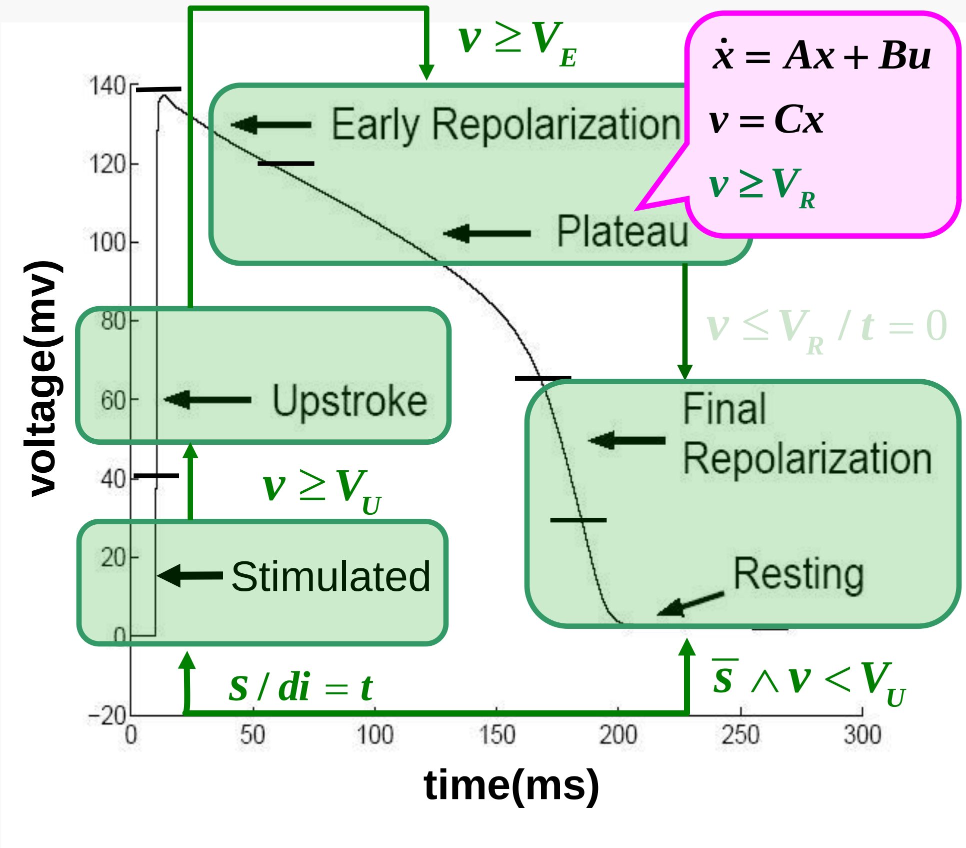
Radu Grosu
SUNY at Stony Brook

How did I get interested in this topic?

Convergence of Theories

- **Hybrid Systems Computation and Control:**
  - **convergence between control and automata theory.**
- **Hybrid Automata: an outcome of this convergence**
  - **modeling formalism for systems exhibiting both discrete and continuous behavior,**
  - **successfully used to model and analyze embedded and biological systems.**

Lack of Common Foundation for HA



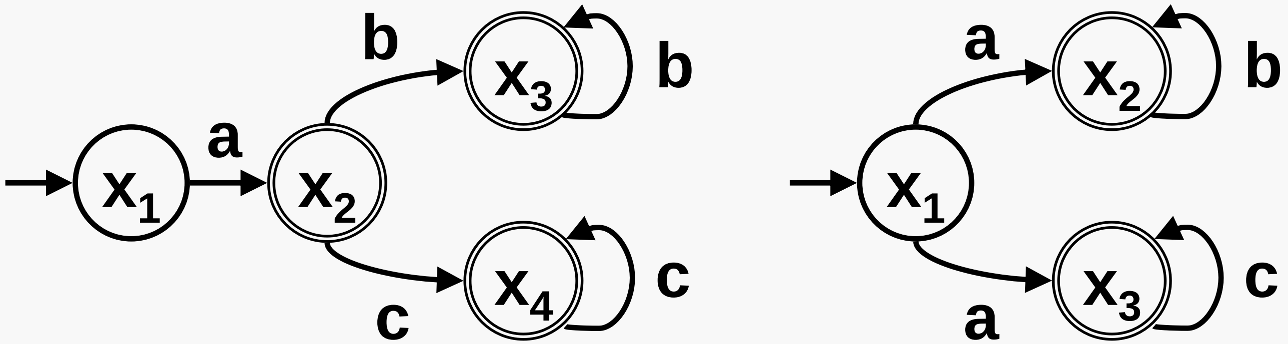
- **Mode dynamics:**
  - **Linear system (LS)**
- **Mode switching:**
  - **Finite automaton (FA)**
- **Different techniques:**
  - **LS reduction**
  - **FA minimization**

- **LS & FA taught separately: No common foundation!**

Main Conjecture

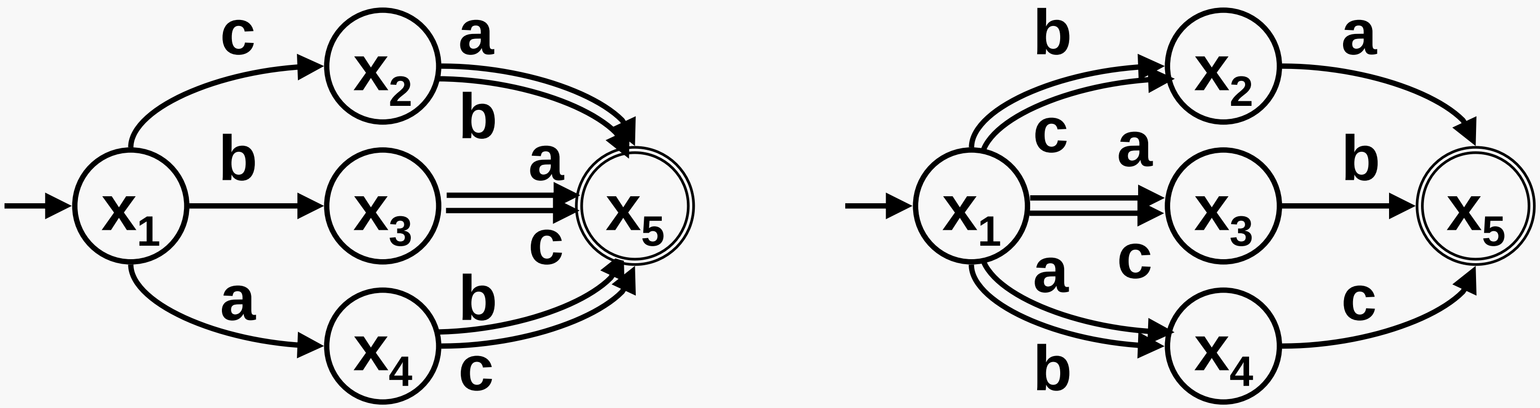
- **Finite automata can be conveniently regarded as time invariant linear systems over semimodules:**
  - **linear systems techniques generalize to automata**
- **Examples of such techniques include:**
  - **linear transformations of automata,**
  - **minimization and determinization of automata as observability and reachability reductions**
  - **Z-transform of automata to compute associated regular expression through Gaussian elimination.**

Minimal DFA are Not Minimal NFA
(Arnold, Dicky and Nivat's Example)



$L = a\ (b^* + c^*)$

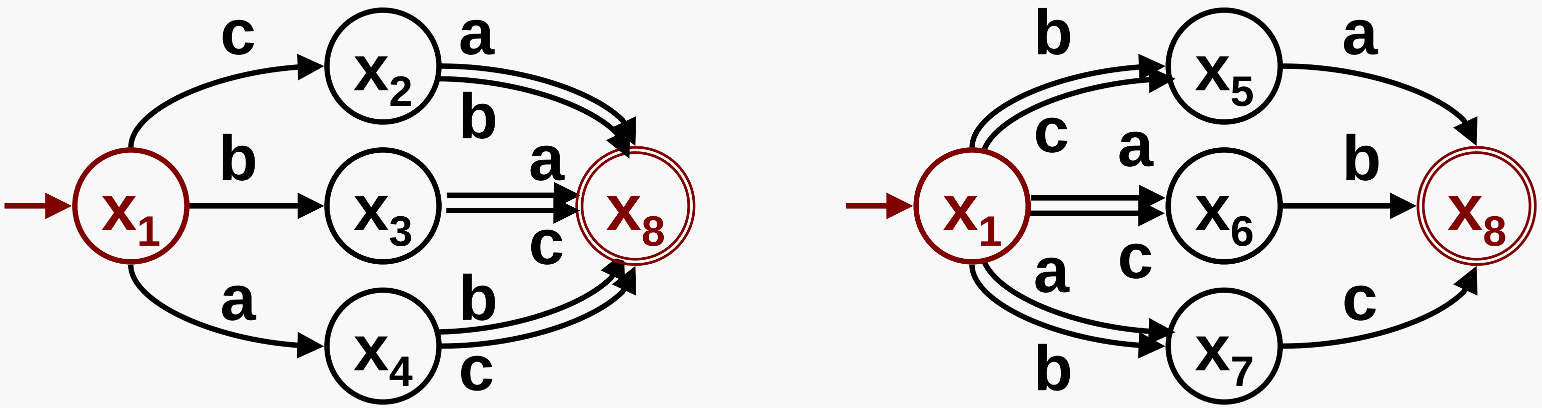
Minimal NFA: How are they Related?
(Arnold, Dicky and Nivat's Example)



$L = ab+ac + ba+bc + ca+cb$

No homomorphism of either automaton onto the other.

Minimal NFA: How are they Related?
(Arnold, Dicky and Nivat's Example)

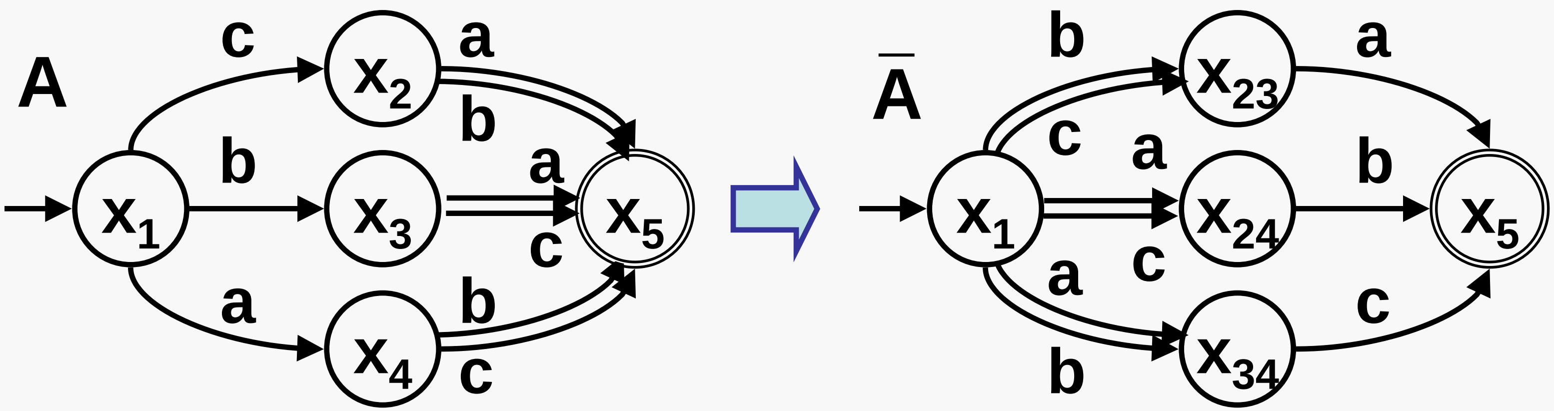


Carrez's solution: Take both in a **terminal NFA.**

Is this the best one can do?

No! One can use use linear (similarity) transformations.

Observability Reduction HSCC'09
(Arnold, Dicky and Nivat's Example)



Define linear transformation $\bar{x}^t = x^t T$:

$T =$

	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$
x_1	1	0	0	0	0
x_2	0	1	1	0	0
x_3	0	1	0	1	0
x_4	0	0	1	1	0
x_5	0	0	0	0	1

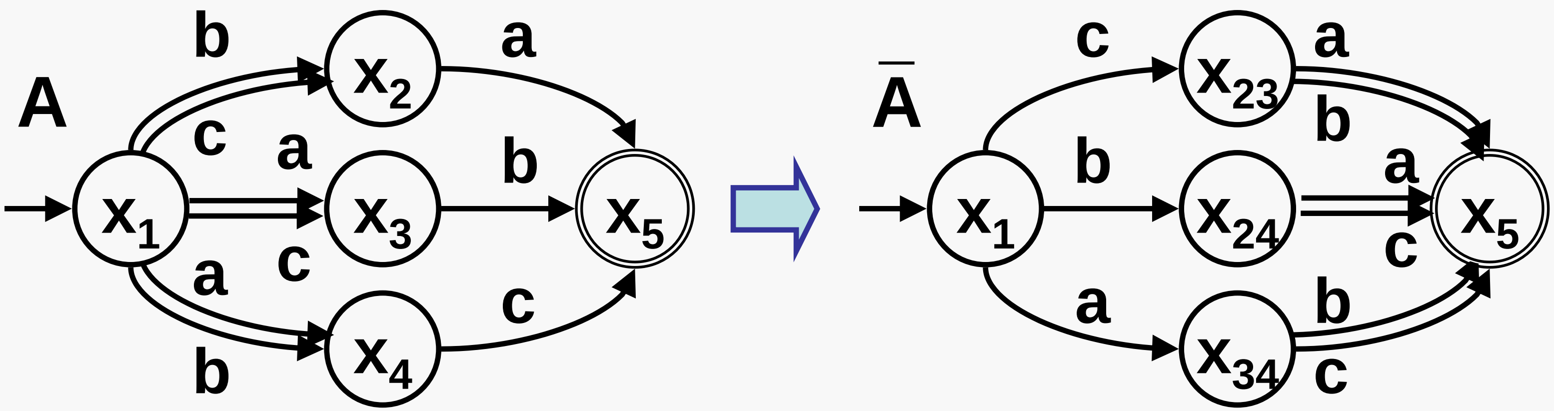
$\bar{A} = [AT]_T$ $(T^{-1}AT)$

$\bar{x}_0^t = x_0^t T$

$\bar{C} = [C]_T$ $(T^{-1}C)$

Reachability Reduction HSCC'09

(Arnold, Dicky and Nivat's Example)



Define linear transformation $\bar{x}^t = x^t T$:

$T =$

	x_1	x_2	x_3	x_4	x_5
x_1	1	0	0	0	0
x_2	0	1	1	0	0
x_3	0	1	0	1	0
x_4	0	0	1	1	0
x_5	0	0	0	0	1

$\bar{A}^t = [A^t T]_T$ $(T^{-1} A^t T)$

$\bar{x}_0^t = x_0^t T$

$\bar{C} = [C]_T$ $(T^{-1} C)$

First improvement of fundamental algorithm in 10 years

The max-flow problem, which is ubiquitous in network analysis, scheduling, and logistics, can now be solved more efficiently than ever.

Larry Hardesty, MIT News Office

September 27, 2010

Graphic: Christine Daniloff

The maximum-flow problem, or max flow, is one of the most basic problems in computer science: First solved during preparations for the Berlin airlift, it's a component of many logistical problems and a staple of introductory courses on algorithms. For decades it was a prominent research subject, with new algorithms that solved it more and more efficiently coming out once or twice a year. But as the problem became better understood, the pace of innovation slowed. Now, however, MIT researchers, together with colleagues at Yale and the University of Southern California, have demonstrated the first improvement of the max-flow algorithm in 10 years.

massachusetts institute of technology

MITnews

search

go

In the branch of mathematics known as linear algebra, a row of a matrix can also be interpreted as a mathematical equation, and the tools of linear algebra enable the simultaneous solution of all the equations embodied by all of a matrix's rows. By repeatedly modifying the numbers in the matrix and re-solving the equations, the researchers effectively evaluate the whole graph at once. This approach, which Kelner will describe at a talk at MIT's Stata Center on Sept. 28, turns out to be more efficient than trying out paths one by one.

Graphic: Christine Daniloff

With the Web, people worldwide can work on distributed tasks. But getting reliable results requires algorithms that specify workflow between people, not transistors.

or max flow ... of the most basic problems in computer science: First solved during preparations for the Berlin airlift, it's a component of many logistical problems and a staple of introductory courses on algorithms. For decades it was a prominent research subject, with new algorithms that solved it more and more efficiently coming out once or twice a year. But as the problem became better understood, the pace of innovation slowed. Now, however, MIT researchers, together with colleagues at Yale and the University of Southern California, have demonstrated the first improvement of the max-flow algorithm in 10 years.

Graphic: Christine Daniloff

Faster Approximation of Maximum Flow in Undirected Graphs" (PDF)

Jonathan Kelner

ARCHIVE: "Unraveling the Matrix"

tags

algorithms

computer science and artificial intelligence laboratory

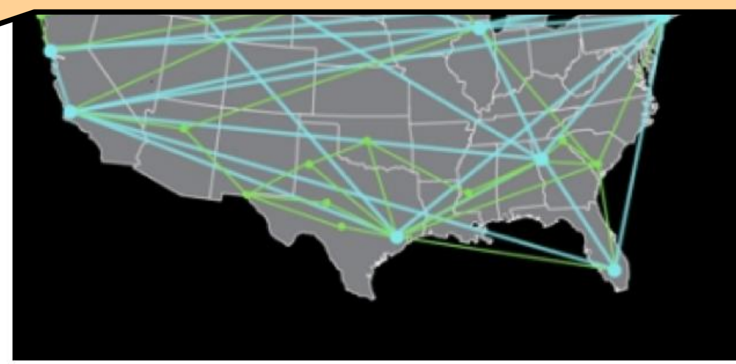
The immediate practicality of the algorithm, however, is not what impresses John Hopcroft, the IBM Professor of Engineering and Applied Mathematics at Cornell and a recipient of the Turing Prize, the highest award in computer science. "My guess is that this particular framework is going to be applicable to a wide range of other problems," Hopcroft says. "It's a fundamentally new technique. When there's a breakthrough of that nature, usually, then, a subdiscipline forms, and in four or five years, a number of results come out."

Graphic: Christine Daniloff

With the Web, people worldwide can work on distributed tasks. But getting reliable results requires algorithms that specify workflow between people, not transistors.

'Build back better'

most basic ... ems in computer science: First solved during preparations for the Berlin airlift, it's a component of many logistical problems and a staple of introductory courses on algorithms. For decades it was a prominent research subject, with new algorithms that solved it more and more efficiently coming out once or twice a year. But as the problem became better understood, the pace of innovation slowed. Now, however, MIT researchers, together with colleagues at Yale and the University of Southern California, have demonstrated the first improvement of the max-flow algorithm in 10 years.



Graphic: Christine Daniloff

... of Maximum Flow in Undirected Graphs" (PDF)

Jonathan Kelner

ARCHIVE: "Unraveling the Matrix"

tags

algorithms

computer science and artificial intelligence laboratory

Observability and minimization

Finite Automata as Linear Systems

- Consider a finite automaton $M = (X, \Sigma, \delta, S, F)$ with:
  - finite set of states X, finite input alphabet Σ,
  - transition relation $\delta \subseteq X \times \Sigma \times X$,
  - starting and final sets of states $S, F \subseteq X$

Finite Automata as Linear Systems

- Consider a finite automaton $M = (X,\Sigma,\delta,S,F)$ with:
  - finite set of states X, finite input alphabet Σ,
  - transition relation $\delta \subseteq X \times \Sigma \times X$,
  - starting and final sets of states $S,F \subseteq X$
- Let X denote row and column indices. Then:
  - δ defines a matrix A,
  - S and F define corresponding vectors

Finite Automata as Linear Systems

- Now define the linear system $L_M = [S, A, C]$:

$$x^t(n+1) = x^t(n)A, \quad x_0 = S$$

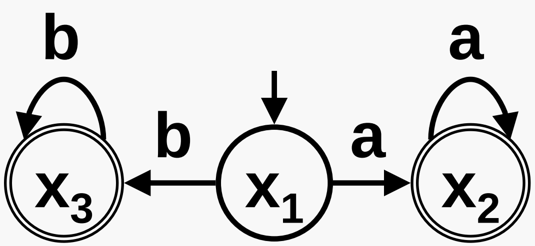
$$y^t(n) = x^t(n)C, \quad C = F$$

Finite Automata as Linear Systems

- Now define the linear system $L_M = [S,A,C]$:

$$x^t(n+1) = x^t(n)A, \quad x_0 = S$$
$$y^t(n) = x^t(n)C, \quad C = F$$

- Example: consider following automaton:



$$A = \begin{bmatrix} 0 & a & b \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

$$x_0 = \begin{bmatrix} \varepsilon \\ 0 \\ 0 \end{bmatrix} \quad C = \begin{bmatrix} 0 \\ \varepsilon \\ \varepsilon \end{bmatrix}$$

Semimodule of Languages

- $\wp(\Sigma^*)$ is an idempotent semiring (quantale):
  - $(\wp(\Sigma^*),+,0)$ is a commutative idempotent monoid (union),
  - $(\wp(\Sigma^*),\times,1)$ is a monoid (concatenation),
  - multiplication distributes over addition,
  - 0 is an annihilator: $0 \times a = 0$

- $(\wp(\Sigma^*))^n$ is a semimodule over scalars in $\wp(\Sigma^*)$:
  - $r(x+y) = rx + ry$, $(r+s)x = rx + sx$, $(rs)x = r(sx)$,
  - $1x = x$, $0x = 0$

- Note: No additive and multiplicative inverses!

Semimodule of Languages

- $\wp(\Sigma^*)$ is an idempotent semiring (quantale):
  - ($\wp(\Sigma^*)$,+,0) is a commutative idempotent monoid (union),
  - ($\wp(\Sigma^*)$,$\times$,1) is a monoid (concatenation),
  - multiplication distributes over addition,
  - 0 is an annihilator: $0 \times a = 0$
- $(\wp(\Sigma^*))^n$ is a semimodule over scalars in $\wp(\Sigma^*)$:
  - r(x+y) = rx + ry, (r+s)x = rx + sx, (rs)x = r(sx),
  - 1x = x, 0x = 0
- **Note:** No additive and multiplicative inverses!

Observability

- Let L = [S,A,C]. Observe its output upto n-1:

$$[y(0)\ y(1)\ \dots\ y(n-1)] = x_0^t\,[C\ AC\ \dots\ A^{n-1}C] = x_0^t O \qquad (1)$$

- If L operates on a vector space:
  - L is observable if: x_0 is uniquely determined by (1),
  - Observability matrix O: has rank n,
  - n-outputs suffice: $A^nC = s_1A^{n-1}C + s_2A^{n-2}C + \dots + s_nC$
- If L operates on a semimodule:
  - L is observable if: x_0 is uniquely determined by (1)

Observability

- Let L = [S,A,C]. Observe its output upto n-1:

$$[y(0)\ y(1)\ \ldots\ y(n\text{-}1)] = x_0^t\,[C\ AC\ \ldots\ A^{n-1}C] = x_0^t O \qquad (1)$$

- If L operates on a vector space:
  - L is observable if: x_0 is uniquely determined by (1),
  - Observability matrix O: has rank n,
  - n-outputs suffice: $A^nC = s_1A^{n-1}C + s_2A^{n-2}C + \ldots + s_nC$
    (Cayley-Hamilton Theorem)

- If L operates on a semimodule:
  - L is observable if: x_0 is uniquely determined by (1)

Observability

- Let L = [S,A,C]. Observe the output upto n-1:

  $$[y(0)\ y(1)\ \dots\ y(n-1)] = x_0^t [C\ AC\ \dots\ A^{n-1}C] = x_0^t O \qquad (1)$$

- If L operates on a vector space:
  - L is observable if: x_0 is uniquely determined by (1),
  - Observability matrix O: has rank n,
  - n-outputs suffice: $A^n C = s_1 A^{n-1} C + s_2 A^{n-2} C + \dots + s_n C$
- If L operates on a semimodule:
  - L is observable if: x_0 is uniquely determined by (1)

The Cayley-Hamilton Theorem

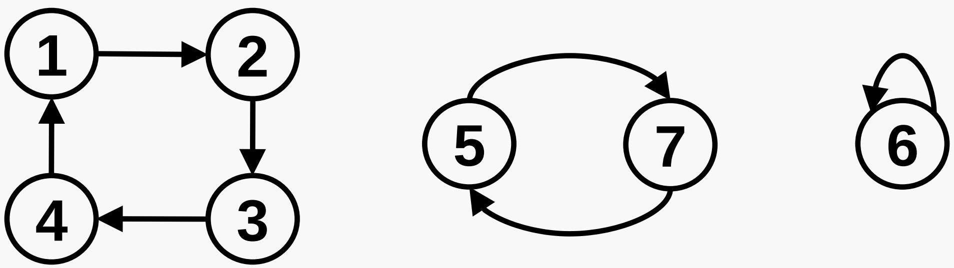
$$(A^n = s_1 A^{n-1} + s_2 A^{n-2} + \ldots + s_n I)$$

Permutations

- **Permutations are bijections of {1,...,n}:**
  - **Example:** π **= {(1,2),(2,3),(3,4),(4,1),(5,7),(6,6),(7,5)}**
- **The graph G(π) of a permutation π:**
  - **G(π) decomposes into: elementary cycles,**
- **The sign of a permutation:**
  - **Pos/Neg: even/odd number of even length cycles,**
  - **P_n^+ / P_n^-: all positive/negative permutations.**

Permutations

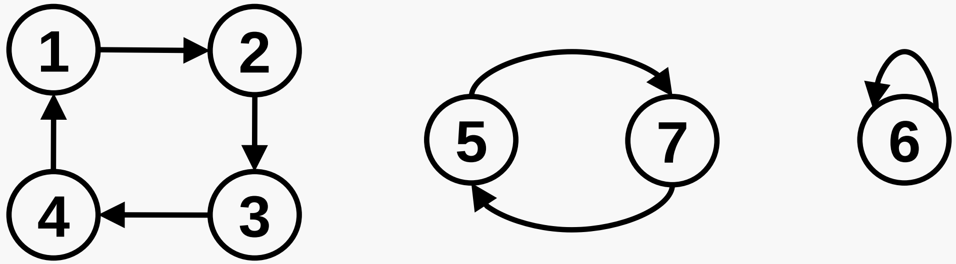
- Permutations are bijections of {1,...,n}:
  - Example: $\pi = \{(1,2),(2,3),(3,4),(4,1),(5,7),(6,6),(7,5)\}$
- The graph $G(\pi)$ of a permutation π:
  - $G(\pi)$ decomposes into: elementary cycles



- The sign of a permutation:
  - Pos/Neg: even/odd number of even length cycles
  - P_n^+ / P_n^-: all positive/negative permutations.

Permutations

- Permutations are bijections of {1,...,n}:
  - Example: π = {(1,2),(2,3),(3,4),(4,1),(5,7),(6,6),(7,5)}
- The graph G(π) of a permutation π:
  - G(π) decomposes into: elementary cycles



- The sign of a permutation π:
  - **Pos/Neg:** even/odd number of even length cycles
  - P_n^+ / P_n^-: all positive/negative permutations

Eigenvalues in Vector Spaces

- **The eigenvalues of a square matrix A:**
  - Eigenvector equation: $x^t A = x^t s$

eigenvector

eigenvalue

- **The characteristic equation of A:**
  - The characteristic polynomial: $cp_A(s) = |sI-A|$
  - The characteristic equation: $cp_A(s) = 0$
- **The determinant of A:**
  - The determinant: $|A| = \sum_{\pi \in P_n^+} \pi(A) - \sum_{\pi \in P_n^-} \pi(A)$,
  - Permutation application: $\pi(A) = \prod_{i=1}^{n} A(i,\pi(i))$

Matrix-Eigenspaces in Vector Spaces

- **The eigenvalues of a square matrix A:**
  - Eigenvector equation: $x^t(sI-A) = 0$

- **The characteristic equation of A:**
  - The characteristic polynomial: $cp_A(s) = |sI-A|$
  - The characteristic equation: $cp_A(s) = 0$

- **The determinant of A:**
  - The determinant: $|A| = \sum_{\pi \in P_n^+} \pi(A) - \sum_{\pi \in P_n^-} \pi(A)$,
  - Permutation application: $\pi(A) = \prod_{i=1}^{n} A(i,\pi(i))$

Matrix-Eigenspaces in Vector Spaces

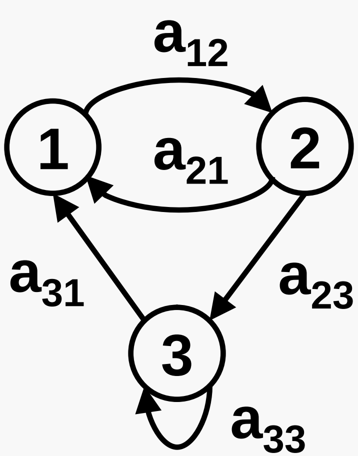
- The eigenvalues of a square matrix A:
  - Eigenvector equation: $x^t(sI-A) = 0$
- The characteristic equation of A:
  - The characteristic polynomial: $cp_A(s) = |sI-A|$
  - The characteristic equation: $cp_A(s) = 0$
- The determinant of A:
  - The determinant: $|A| = \sum_{\pi \in P_n^+} \pi(A) - \sum_{\pi \in P_n^-} \pi(A)$,
  - Permutation application: $\pi(A) = \prod_{i=1}^{n} A(i,\pi(i))$

Matrix-Eigenspaces in Vector Spaces

- The eigenvalues of a square matrix A:
  - Eigenvector equation: $x^t(sI-A) = 0$
- The characteristic equation of A:
  - The characteristic polynomial: $cp_A(s) = |sI-A|$
  - The characteristic equation: $cp_A(s) = 0$
- The determinant of A:
  - The determinant: $|A| = \sum_{\pi \in P_n^+} \pi(A) - \sum_{\pi \in P_n^-} \pi(A)$,
  - Weight of a permutation: $\pi(A) = \prod_{i=1}^{n} A(i,\pi(i))$

The Cayley-Hamilton Theorem (CHT)

- **A satisfies its characteristic equation:** $cp_A(A) = 0$



$$A = \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \qquad sI-A = \begin{bmatrix} s & -a_{12} & 0 \\ -a_{21} & s & -a_{23} \\ -a_{31} & 0 & s-a_{33} \end{bmatrix}$$

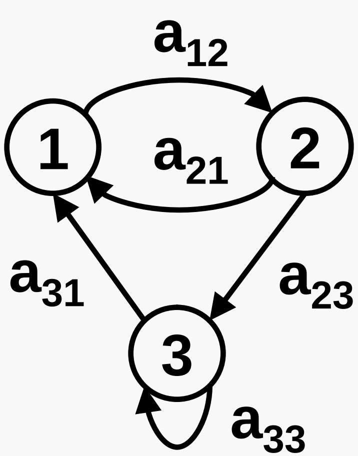
$$|sI-A| = s^3 - a_{33}s^2 - a_{12}a_{21}s + a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31} = 0$$

$$s^3 + a_{12}a_{21}a_{33} = a_{33}s^2 + a_{12}a_{21}s + a_{12}a_{23}a_{31}$$

$$A^3 + a_{12}a_{21}a_{33}I = a_{33}A^2 + a_{12}a_{21}A + a_{12}a_{23}a_{31}I$$

The Cayley-Hamilton Theorem (CHT)

- **A satisfies its characteristic equation:** $cp_A(A) = 0$



$$A = \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \qquad sI-A = \begin{bmatrix} s & -a_{12} & 0 \\ -a_{21} & s & -a_{23} \\ -a_{31} & 0 & s-a_{33} \end{bmatrix}$$

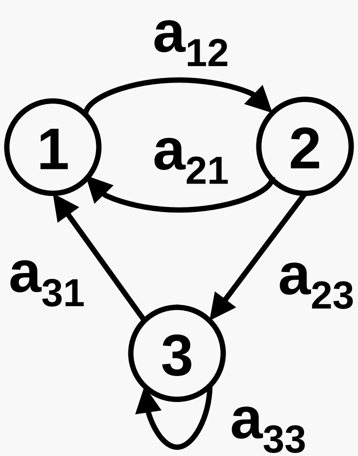
$$|sI-A| = s^3 - a_{33}s^2 - a_{12}a_{21}s + a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31} = 0$$

$$s^3 + a_{12}a_{21}a_{33} = a_{33}s^2 + a_{12}a_{21}s + a_{12}a_{23}a_{31}$$

$$A^3 + a_{12}a_{21}a_{33}I = a_{33}A^2 + a_{12}a_{21}A + a_{12}a_{23}a_{31}I$$

The Cayley-Hamilton Theorem (CHT)

- A satisfies its characteristic equation: $cp_A(A) = 0$



$$A = \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \qquad sI-A = \begin{bmatrix} s & -a_{12} & 0 \\ -a_{21} & s & -a_{23} \\ -a_{31} & 0 & s-a_{33} \end{bmatrix}$$

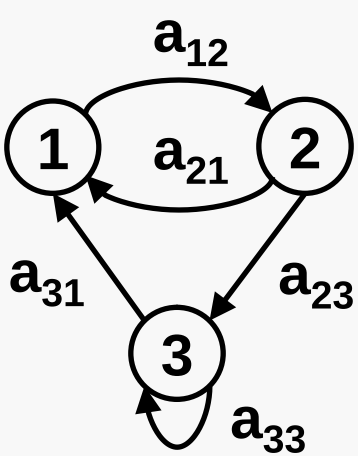
$$|sI-A| = s^3 - a_{33}s^2 - a_{12}a_{21}s + a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31} = 0$$

$$s^3 + a_{12}a_{21}a_{33} = a_{33}s^2 + a_{12}a_{21}s + a_{12}a_{23}a_{31}$$

$$A^3 + a_{12}a_{21}a_{33}I = a_{33}A^2 + a_{12}a_{21}A + a_{12}a_{23}a_{31}I$$

The Cayley-Hamilton Theorem (CHT)

- **A satisfies its characteristic equation:** $cp_A(A) = 0$



$$A = \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \qquad sI\text{-}A = \begin{bmatrix} s & -a_{12} & 0 \\ -a_{21} & s & -a_{23} \\ -a_{31} & 0 & s-a_{33} \end{bmatrix}$$

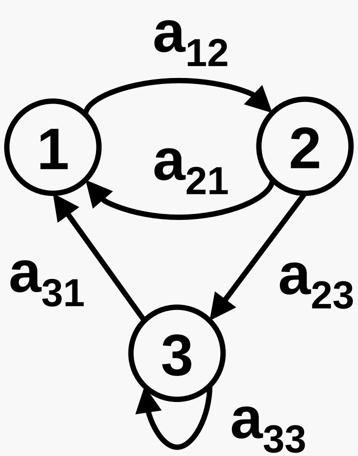
$$|sI-A| = s^3 - a_{33}s^2 - a_{12}a_{21}s + a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31} = 0$$

$$s^3 + a_{12}a_{21}a_{33} = a_{33}s^2 + a_{12}a_{21}s + a_{12}a_{23}a_{31}$$

$$A^3 + a_{12}a_{21}a_{33}I = a_{33}A^2 + a_{12}a_{21}A + a_{12}a_{23}a_{31}I$$

The Cayley-Hamilton Theorem (CHT)

- A satisfies its characteristic equation: $cp_A(A) = 0$



$$A = \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \qquad sI-A = \begin{bmatrix} s & -a_{12} & 0 \\ -a_{21} & s & -a_{23} \\ -a_{31} & 0 & s-a_{33} \end{bmatrix}$$

$$|sI-A| = s^3 - a_{33}s^2 - a_{12}a_{21}s + a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31} = 0$$

$$s^3 + a_{12}a_{21}a_{33} = a_{33}s^2 + a_{12}a_{21}s + a_{12}a_{23}a_{31}$$

$$A^3 + a_{12}a_{21}a_{33}I = a_{33}A^2 + a_{12}a_{21}A + a_{12}a_{23}a_{31}I$$

cycle cycle cycle cycle cycle

The Cayley-Hamilton Theorem (CHT)

- A satisfies its characteristic equation: $cp_A(A) = 0$
- Implicit assumptions in CHT:
  - Subtraction is available
  - Multiplication is commutative
- Does CHT hold in semirings?
  - Subtraction not indispensible (Rutherford, Straubing)
  - Commutativity still problematic

The Cayley-Hamilton Theorem (CHT)

- A satisfies its characteristic equation: $cp_A(A) = 0$
- Implicit assumptions in CHT:
  - Subtraction is available
  - Multiplication is commutative
- **Does CHT hold in semirings?**
  - Subtraction not indispensible (Rutherford, Straubing)
  - Commutativity problematic

CHT in Commutative Semirings
(Straubing's Proof)

- **Lift original semiring to the semiring of paths:**
  - **Matrix A is lifted to a matrix G_A of paths π**

$$A = \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \Rightarrow G_A = \begin{bmatrix} 0 & (1,2) & 0 \\ (2,1) & 0 & (2,3) \\ (3,1) & 0 & (3,3) \end{bmatrix}$$

CHT in Commutative Semirings
(Straubing's Proof)

- **Lift original semiring to the semiring of paths:**
  - **- Matrix A is lifted to a matrix G_A of paths π**
  - **- Permutation cycles σ lifted cyclic paths π_σ**

σ = {(1,2),(2,1)} ⇒ **π_σ = (1,2)(2,1)**

CHT in Commutative Semirings
(Straubing's Proof)

- Lift original semiring to the semiring of paths:
  - Matrix A is lifted to a matrix G_A of paths π
  - Permutation cycles lifted cyclic paths π_σ
- **Prove CHT in the semiring of paths:**

$$\sum_{q=0}^{n} \sum_{\sigma \in P_q^{+}} \pi_\sigma G_A^{n-q} = \sum_{q=0}^{n} \sum_{\sigma \in P_q^{-}} \pi_\sigma G_A^{n-q}$$

(CHT holds?)

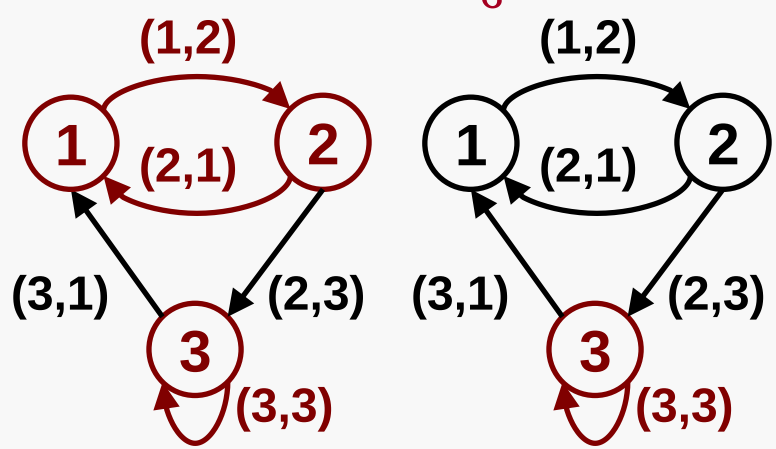
CHT in Commutative Semirings
(Straubing's Proof)

- Lift original semiring to the semiring of paths:
  - Matrix A is lifted to a matrix G_A of paths π
  - Permutation cycles lifted cyclic paths π_σ

- **Prove CHT in the semiring of paths:**
  - **Show bijection** between pos/neg products $\pi_\sigma \pi$

$$\sum_{\sigma \in P_3^+} \pi_\sigma G_A^0 = \sum_{\sigma \in P_1^-} \pi_\sigma G_A^2$$

(3,3)(1,2)(2,1) ⇔ (3,3)(1,2)(2,1)



CHT in Commutative Semirings
(Straubing's Proof)

- Lift original semiring to the semiring of paths:
  - Matrix A is lifted to a matrix G_A of paths π
  - Permutation cycles lifted cyclic paths π_σ
- Prove CHT in the semiring of paths:
  - Show bijection between pos/neg products $\pi_\sigma \pi$
- Port results back to the original semiring:
  - Apply products: $\pi_\sigma \pi(A)$
  - Path application: $(\pi_1 \ldots \pi_n)(A) = A(\pi_1) \ldots A(\pi_n)$

CHT in Idempotent Semirings

- Lift original semiring to the semiring of paths:
  - Matrix A: order in paths π important
  - Permutation cycles: rotations are distinct

CHT in Idempotent Semirings

- **Lift original semiring to the semiring of paths:**
  - **Matrix A:** order in paths π important
  - **Permutation cycles:** rotations are distinct

$$\sigma = \{(1,2),(2,1)\} \Rightarrow \Pi_\sigma = \begin{bmatrix} (1,2)(2,1) & 0 & 0 \\ 0 & (2,1)(1,2) & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

CHT in Idempotent Semirings

- Lift original semiring to the semiring of paths:
  - Matrix A: order in paths π important
  - Permutation cycles: rotations are distinct
- Prove CHT in the semiring of paths:
  - Products $\pi_\sigma G^{n-|\sigma|}$: cycles to be properly inserted

CHT in Idempotent Semirings

- Lift original semiring to the semiring of paths:
  - Matrix A: order in paths π important
  - Permutation cycles: rotations are distinct
- Prove CHT in the semiring of paths:
  - Products $\pi_\sigma G^{n-|\sigma|}$: cycles to be properly inserted

$$\Pi_\sigma * G^{n-|\sigma|} = \Pi_\sigma G^{n-|\sigma|} + G\Pi_\sigma G^{n-|\sigma|-1} + \ldots + G^{n-|\sigma|}\Pi_\sigma$$

CHT in Idempotent Semirings

- Lift original semiring to the semiring of paths:
  - Matrix A: order in paths π important
  - Permutation cycles: rotations are distinct
- Prove CHT in the semiring of paths:
  - Products $\pi_\sigma G^{n-|\sigma|}$: cycles to be properly inserted
- Port results back to the original semiring:
  - Apply products: $\Pi_\sigma G^{n-|\sigma|}(A)$

CHT in Idempotent Semirings

- **Theorem:** $$G^n = \sum_{q=1}^{n} \sum_{\sigma \in P_q^{-}} \Pi_\sigma * G_A^{n-|\sigma|}$$

Proof:

LHS $\subseteq$ RHS: Let $\pi \in$ LHS

- Pidgeon-hole: π has at least one cycle π_σ in s
- Structural: π_σ is a simple cycle of length k
- Remove π_σ in π: $\pi[s/\pi_\sigma]$ is in $G^{n-|\sigma|}$
- Shuffle-product: $\Pi_\sigma * G^{n-|\sigma|}$ reinserts π_σ

RHS $\subseteq$ LHS: Let $\pi \in$ RHS

- No wrong path: The shuffle is sound
- Idempotence: Takes care of multiple copies

CHT in Idempotent Semirings

- Theorem: $G^n = \sum_{q=1}^{n} \sum_{\sigma \in P_q^-}^{n} \Pi_\sigma * G_A^{n-|\sigma|}$

Proof:

LHS $\subseteq$ RHS: Let $\pi \in$ LHS

- Pidgeon-hole: π has at least one cycle π_σ in s
- Structural: π_σ is also a simple cycle
- Remove π_σ in π: $\pi[s/\pi_\sigma]$ is in $G^{n-|\sigma|}$
- Shuffle-product: $\Pi_\sigma * G^{n-|\sigma|}$ reinserts π_σ

RHS $\subseteq$ LHS: Let $\pi \in$ RHS

- No wrong path: The shuffle is sound
- Idempotence: Takes care of multiple copies

CHT in Idempotent Semirings

- **Theorem:** $G^n = \sum_{q=1}^{n} \sum_{\sigma \in P_q^-}^{n} \Pi_\sigma * G_A^{n-|\sigma|}$

Proof:

LHS $\subseteq$ RHS: Let $\pi \in$ LHS

- **Pidgeon-hole:** π has at least one cycle π_σ in s
- **Structural:** π_σ is also a simple cycle
- **Remove π_σ in π:** $\pi[s/\pi_\sigma]$ is in $G^{n-|\sigma|}$
- **Shuffle-product:** $\Pi_\sigma * G^{n-|\sigma|}$ reinserts π_σ

RHS $\subseteq$ LHS: Let $\pi \in$ RHS

- **No wrong path:** The shuffle is sound
- **Idempotence:** Takes care of multiple copies

CHT in Idempotent Semirings

- **Define:** $\bar{\Pi}_\sigma(i,i) = \begin{cases} \sigma & \text{if } \Pi_\sigma(i,i) = 0 \\ 0 & \text{if } \Pi_\sigma(i,i) = \sigma \end{cases}$

- **Theorem: classic CHT can be derived by using:**

  - $\sigma\, G^{n-|\sigma|} = \Pi_\sigma * G_\sigma^{n-|\sigma|} + \bar{\Pi}_\sigma * G_{\bar{\sigma}}^{n-|\sigma|}$

  - application of CHT to $G_\sigma^{n-|\sigma|}$ and $G_\sigma^{n-|\sigma|}$

- **Matrix CHT: can be regarded as a constructive version of the pumping lemma.**

CHT in Idempotent Semirings

- Define: $\bar{\Pi}_\sigma(i,i) = \begin{cases} \sigma & \text{if } \Pi_\sigma(i,i) = 0 \\ 0 & \text{if } \Pi_\sigma(i,i) = \sigma \end{cases}$

- Theorem: classic CHT can be derived by using:

  - $\sigma\, G^{n-|\sigma|} = \Pi_\sigma * G_\sigma^{n-|\sigma|} + \bar{\Pi}_\sigma * G_{\bar{\sigma}}^{n-|\sigma|}$
  - application of CHT to $G_\sigma^{n-|\sigma|}$ and $G_\sigma^{n-|\sigma|}$

- Matrix CHT: can be regarded as a constructive version of the pumping lemma.

CHT in Idempotent Semirings

- Define: $\overline{\Pi}_\sigma(i,i) = \begin{cases} \sigma \text{ if } \Pi_\sigma(i,i) = 0 \\ 0 \text{ if } \Pi_\sigma(i,i) = \sigma \end{cases}$

- Theorem: classic CHT can be derived by using:

  - $\sigma\, G^{n-|\sigma|} = \Pi_\sigma * G_\sigma^{n-|\sigma|} + \overline{\Pi}_\sigma * G_{\bar{\sigma}}^{n-|\sigma|}$
  - application of CHT to $G_\sigma^{n-|\sigma|}$ and $G_\sigma^{n-|\sigma|}$

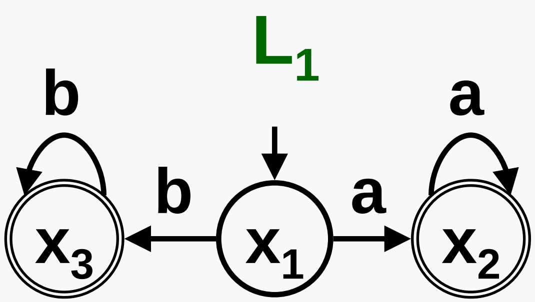
- Matrix CHT: can be regarded as a constructive version of the pumping lemma.

Finite Automata as Linear Systems

- Now define the linear system $L_M = [S,A,C]$:

$$x^t(n+1) = x^t(n)A, \quad x_0 = S(\varepsilon)\varepsilon$$
$$y^t(n) = x^t(n)C, \quad C = F(\varepsilon)\varepsilon$$

- Example: consider following automaton:



$$A = A(a)a + A(b)b$$

$$A(a) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad x_0(\varepsilon) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$A(b) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad C(\varepsilon) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Observability

- Let L = [S,A,C] be an n-state automaton. It's output:

$$[y(0)\ y(1)\ \dots\ y(n\text{-}1)] = x_0^t[C\ AC\ \dots\ A^{n-1}C] = x_0^t O \quad (1)$$

L is observable if x_0 is uniquely determined by (1).

- **Example:** the observability matrix O of L_1 is:

O =

A^nC	ε	a	b	aa	ab	ba	bb
x_1	0	1	1	1	0	0	1
x_2	1	1	0	1	0	0	0
x_3	1	0	1	0	0	0	1