Elias Rojas Collins[1] Chris Köcher[2] Georg Zetzsche[2]

[1] Massachusetts Institute of Technology, Cambridge, USA
[2] Max Planck Institute for Software Systems, Kaiserslautern, Germany

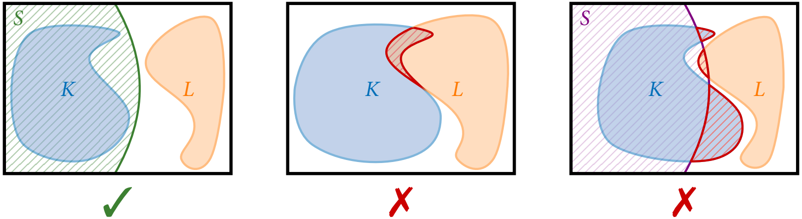
August 28, 2025

Separability Problem

Separability Problem (of two sets from a class $\mathcal{C}$ by a set from a class $\mathcal{D}$)

Input: Two sets $K, L \in \mathcal{C}$.

Question: Is there a set $S \in \mathcal{D}$ such that $K \subseteq S$ and $L \cap S = \emptyset$?



- Regular separability: class $\mathcal{D}$ is the class of regular languages
- Certifies safety of programs:
  - K is the set of reachable configurations in the program
  - L is the set of undesired configurations in the program

Historical Context

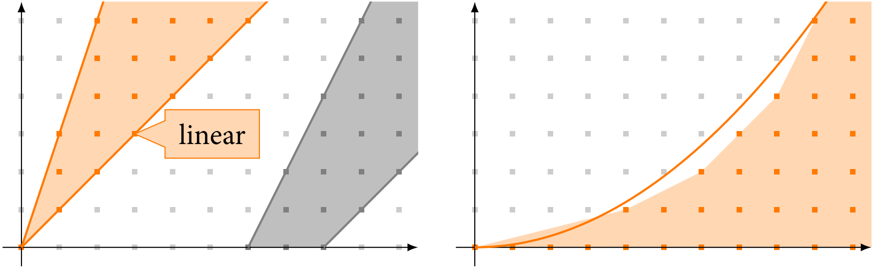
- Regular separability is undecidable for context-free languages [Szymanski & Williams 1976, Hunt III 1982]
- Regular separability is decidable for:
  1. languages of Integer Vector Addition Systems with States ($\mathbb{Z}$-VASS) / Parikh automata [Clemente et al @ STACS 2017]
     - Complexity: elementary, coNP-hard
  2. coverability languages of VASS [Czerwiński et al. @ CONCUR 2018]
     - Complexity: EXPSPACE-complete
  3. (reachability) languages of VASS [Keskin & Meyer @ LICS 2024]
     - Complexity: Ackermann-complete

Semilinear Sets

Definition

A set $S \subseteq \mathbb{N}^d$ is

- linear if there is a vector $\vec{u} \in \mathbb{N}^n$ and a finite set $P \subseteq \mathbb{N}^n$ with $S = \vec{u} + P^*$.
- hyperlinear (or hybrid linear) if there are finite sets $U, P \subseteq \mathbb{N}^n$ with $S = U + P^*$.
- semilinear if it is a finite union of (hyper-)linear sets.

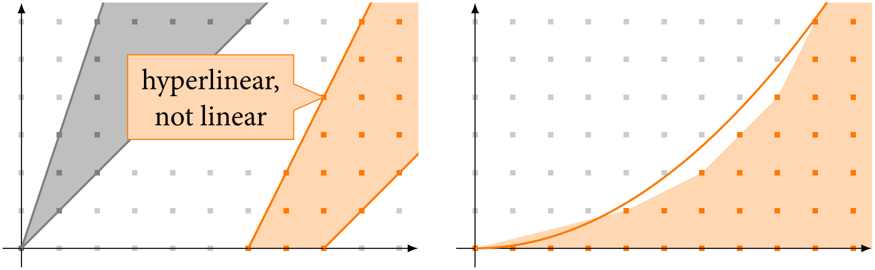


Semilinear Sets

Definition

A set $S \subseteq \mathbb{N}^d$ is

- linear if there is a vector $\vec{u} \in \mathbb{N}^n$ and a finite set $P \subseteq \mathbb{N}^n$ with $S = \vec{u} + P^*$.
- hyperlinear (or hybrid linear) if there are finite sets $U, P \subseteq \mathbb{N}^n$ with $S = U + P^*$.
- semilinear if it is a finite union of (hyper-)linear sets.

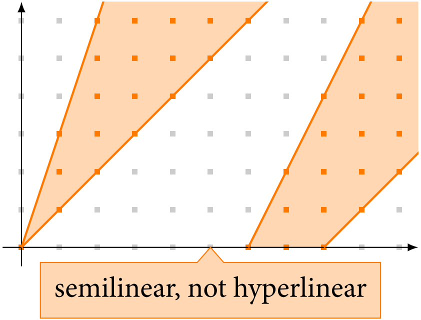


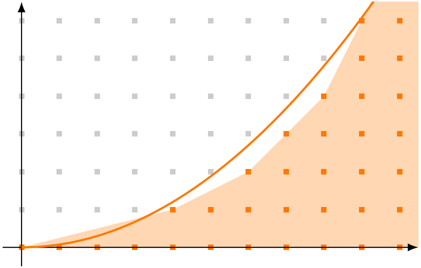
Semilinear Sets

Definition

A set $S \subseteq \mathbb{N}^d$ is

- linear if there is a vector $\vec{u} \in \mathbb{N}^n$ and a finite set $P \subseteq \mathbb{N}^n$ with $S = \vec{u} + P^*$.
- hyperlinear (or hybrid linear) if there are finite sets $U, P \subseteq \mathbb{N}^n$ with $S = U + P^*$.
- semilinear if it is a finite union of (hyper-)linear sets.

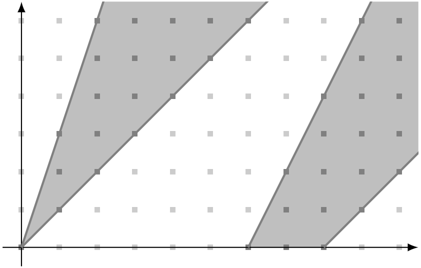


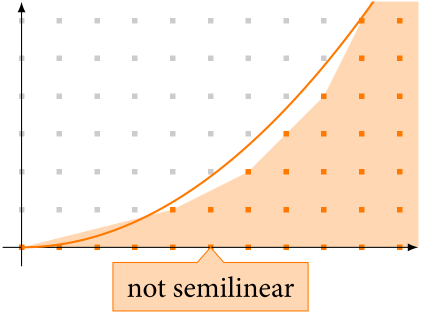


Definition

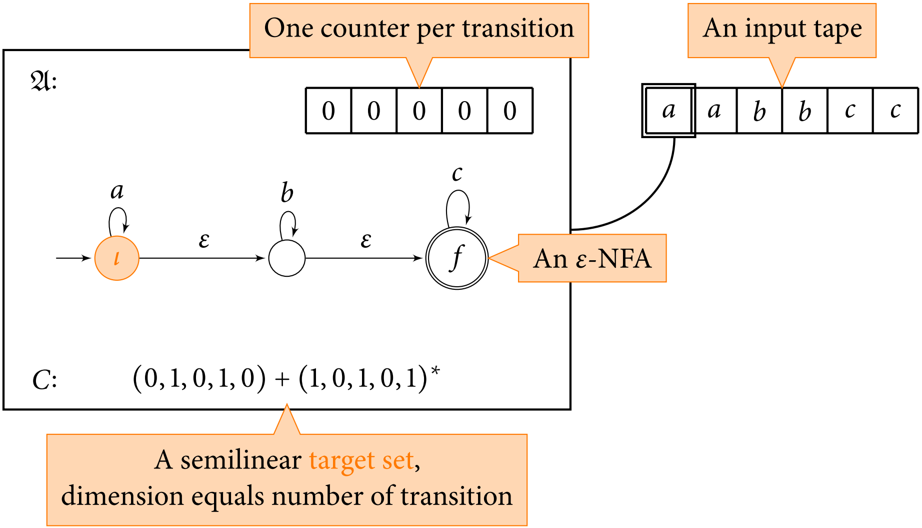
A set $S \subseteq \mathbb{N}^d$ is

- linear if there is a vector $\vec{u} \in \mathbb{N}^n$ and a finite set $P \subseteq \mathbb{N}^n$ with $S = \vec{u} + P^*$.
- hyperlinear (or hybrid linear) if there are finite sets $U, P \subseteq \mathbb{N}^n$ with $S = U + P^*$.
- semilinear if it is a finite union of (hyper-)linear sets.

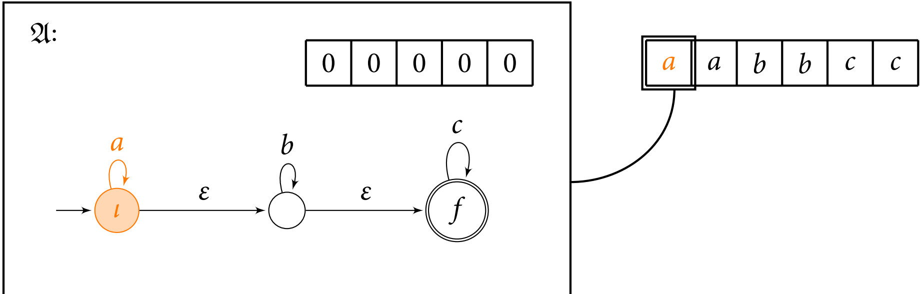




Parikh Automata

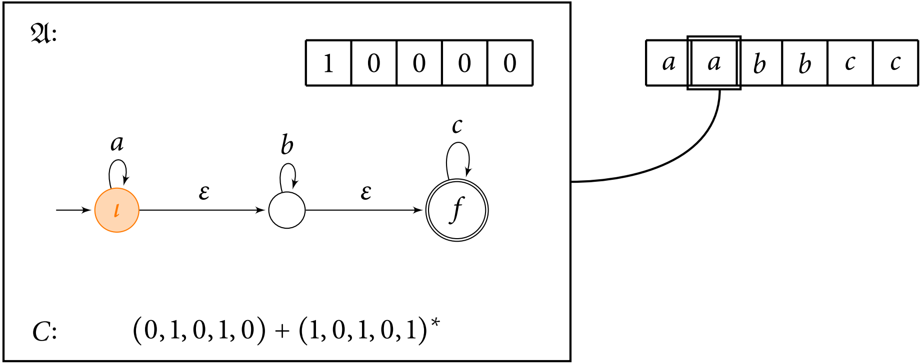


Parikh Automata

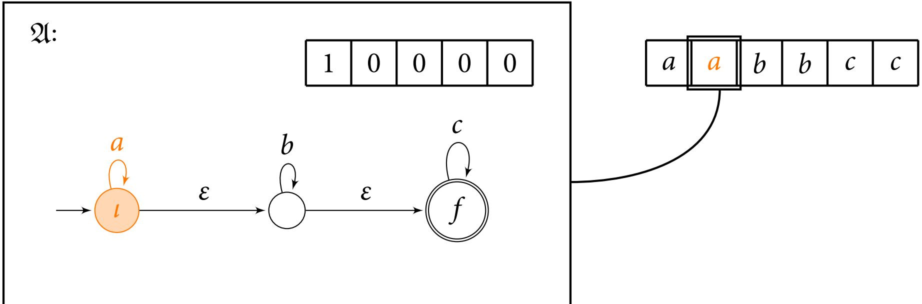


C: $(0,1,0,1,0) + (1,0,1,0,1)^*$

Parikh Automata

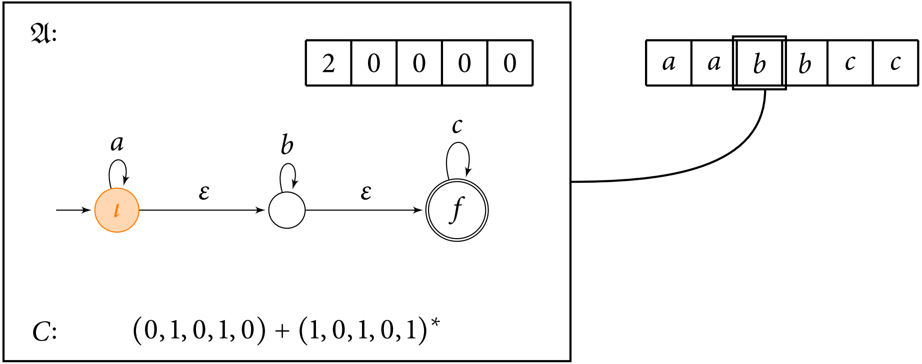


Parikh Automata

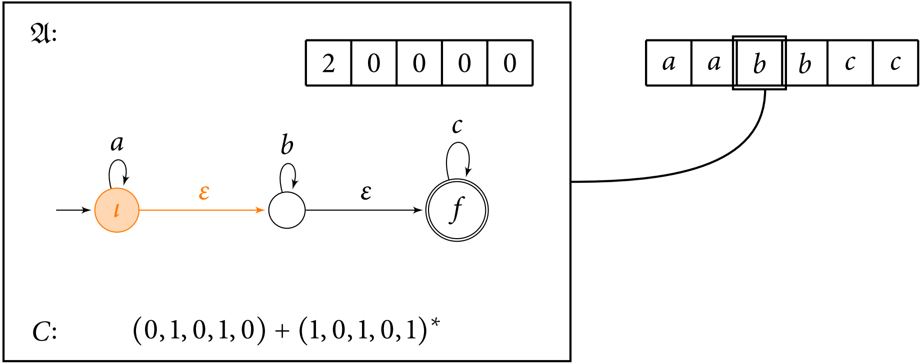


C: $(0,1,0,1,0)+(1,0,1,0,1)^*$

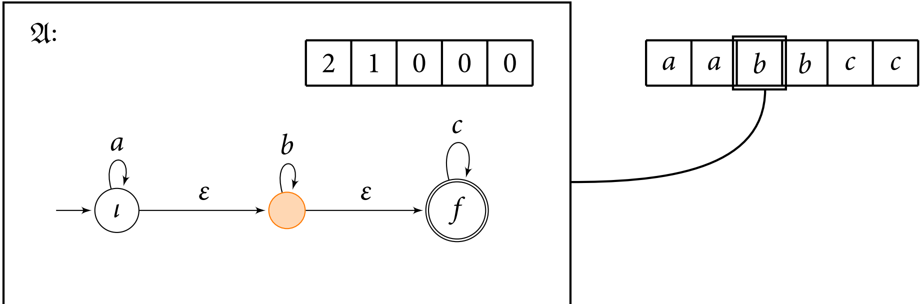
Parikh Automata



Parikh Automata

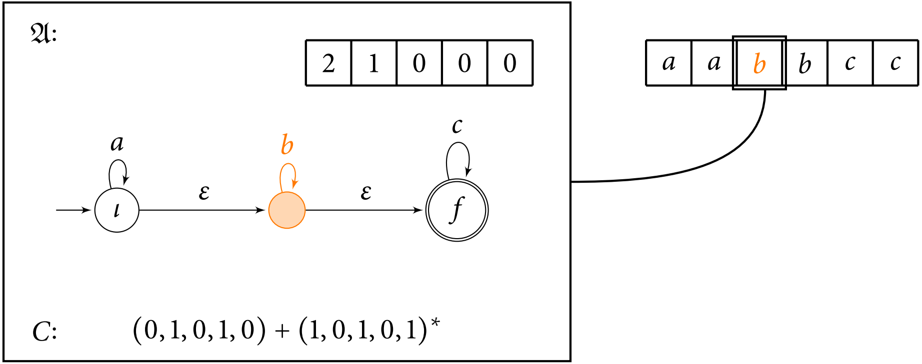


Parikh Automata

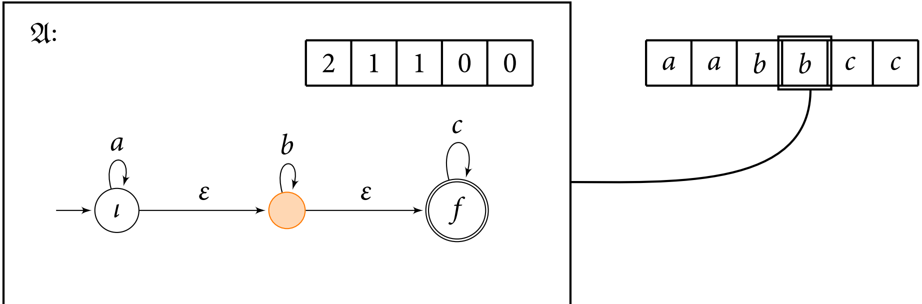


C: $(0,1,0,1,0) + (1,0,1,0,1)^*$

Parikh Automata

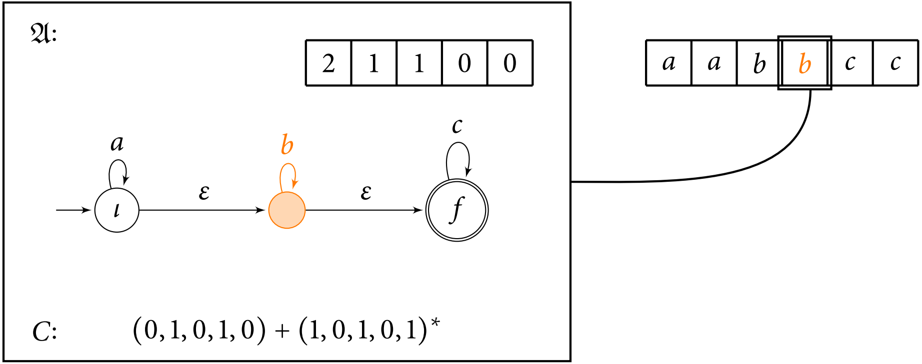


Parikh Automata

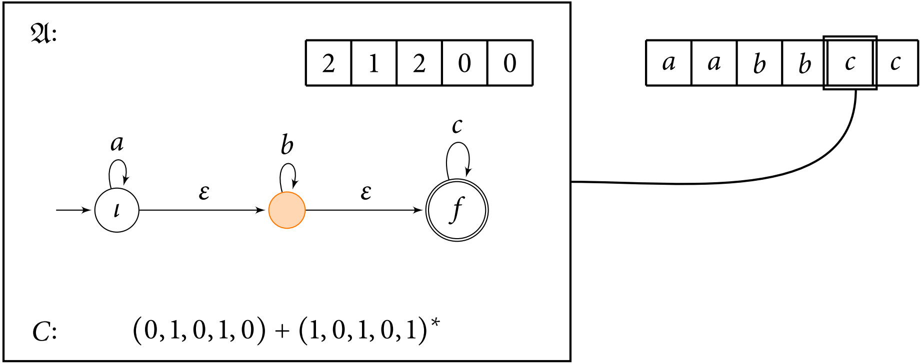


C: $(0,1,0,1,0) + (1,0,1,0,1)^*$

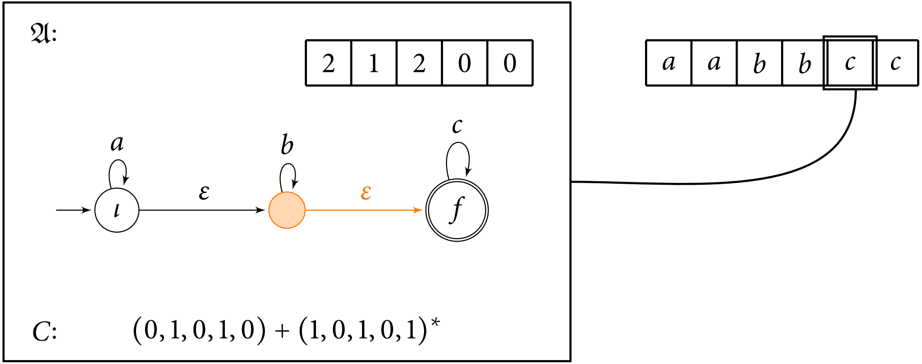
Parikh Automata



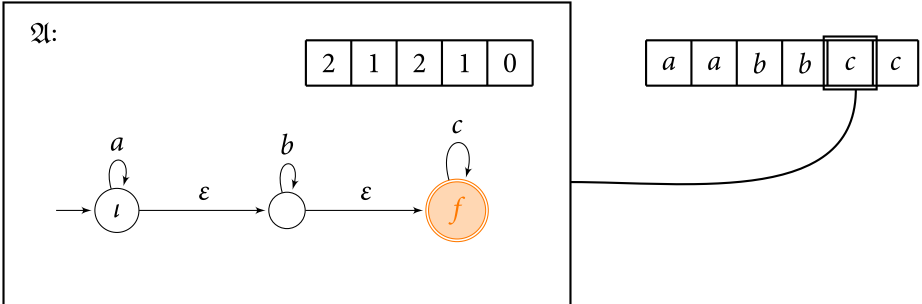
Parikh Automata



Parikh Automata

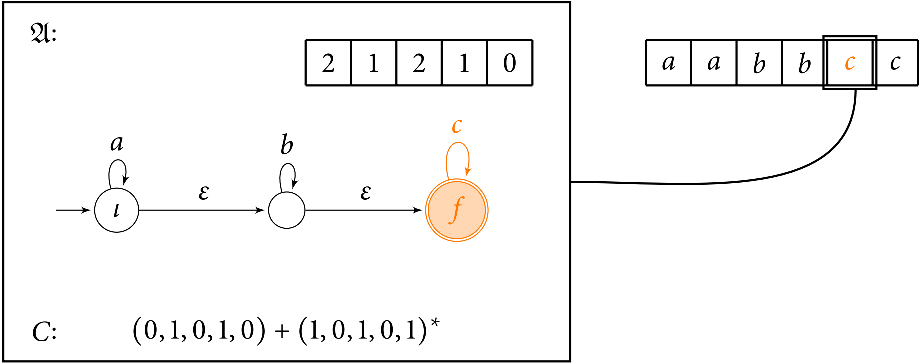


Parikh Automata

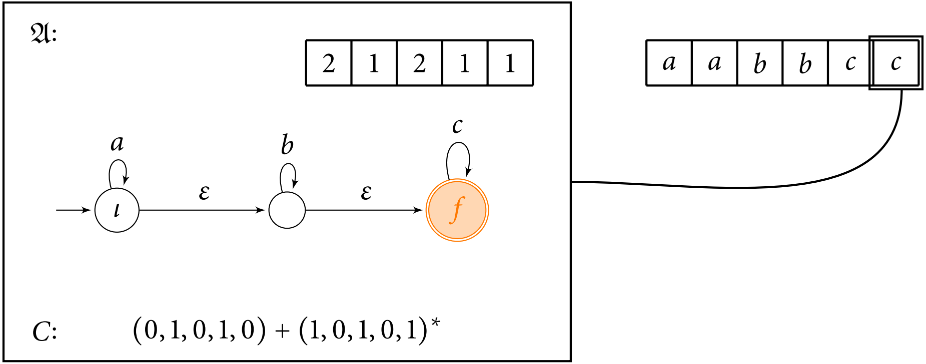


C: $(0, 1, 0, 1, 0) + (1, 0, 1, 0, 1)^*$

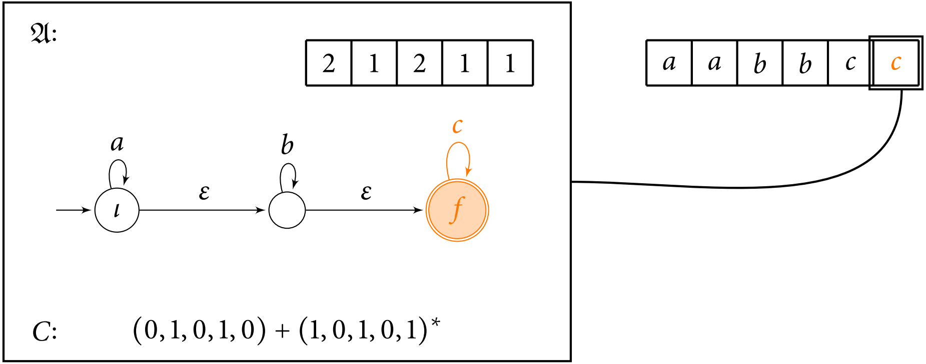
Parikh Automata



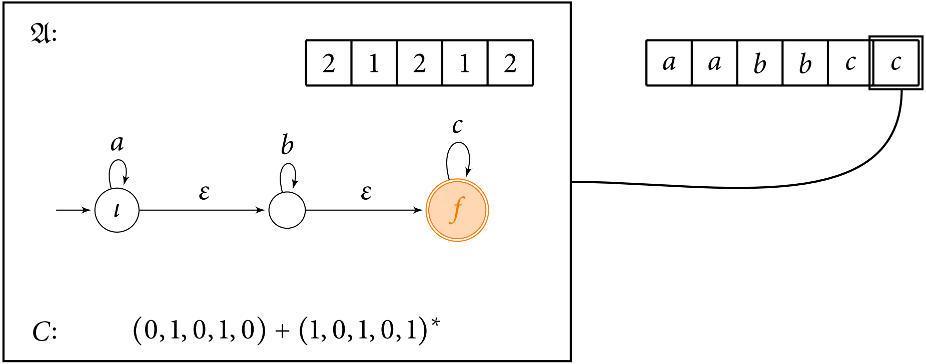
Parikh Automata

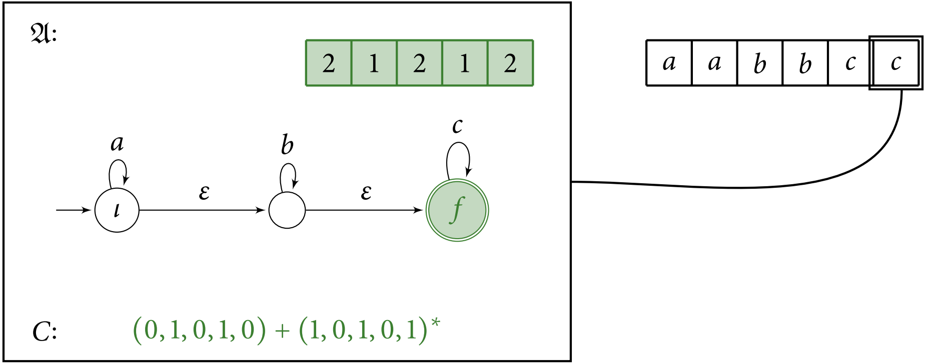


Parikh Automata



Parikh Automata




𝔄:
2 1 2 1 2
a
b
c
ε
ε
ι
f
C:
(0, 1, 0, 1, 0) + (1, 0, 1, 0, 1)*
a a b b c c

Results

Main Theorem

1. The regular separability problem for Parikh automata is coNP-complete.
2. The recognizable separability problem for semilinear sets is coNP-complete.

- We first show that recognizable inseparability of semilinear sets is in NP.

Observation

$R = \bigcup_{1 \leq i \leq m} U_i + P_i^*$ and $S = \bigcup_{1 \leq j \leq n} V_j + Q_j^*$ are inseparable iff there are i and j such that $U_i + P_i^*$ and $V_j + Q_j^*$ are inseparable.

- We guess these i and j in a first step. From now on, assume that R and S are hyperlinear.

Main Theorem

1. The regular separability problem for Parikh automata is coNP-complete.
2. The recognizable separability problem for semilinear sets is coNP-complete.

$S \subseteq \mathbb{M}$ is recognizable in the monoid $\mathbb{M}$
if there is a finite monoid $\mathbb{F}$ and a morphism $f\colon \mathbb{M} \to \mathbb{F}$
with $S = f^{-1}(f(S))$.

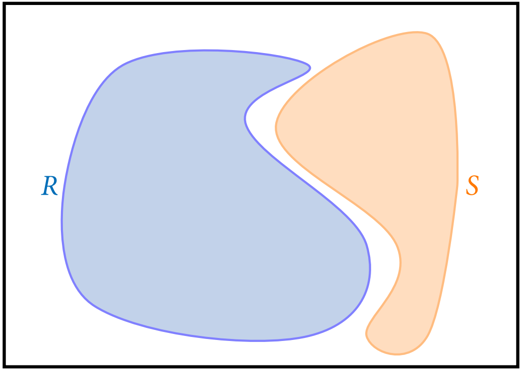
f semilinear sets is in NP.

$R = \bigcup_{1 \leq i \leq m} U_i + P_i^*$ and $S = \bigcup_{1 \leq j \leq n} V_j + Q_j^*$ are inseparable iff there are i and j such that $U_i + P_i^*$ and $V_j + Q_j^*$ are inseparable.

- We guess these i and j in a first step. From now on, assume that R and S are hyperlinear.

A First Idea

- Assume S is bounded at coordinate j, i.e., there is an $M \in \mathbb{N}$ such that $\vec{v}[j] \leq M$ for all $\vec{v} \in S$.

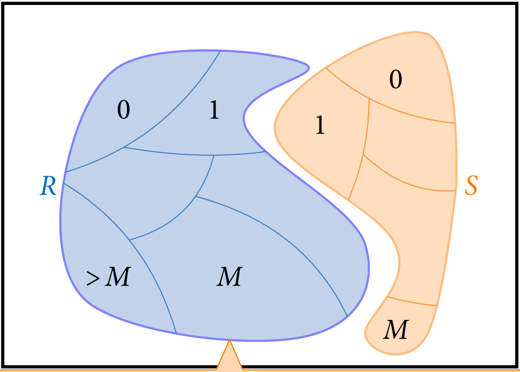


- Problem: Constructing R' and S' is expensive!

[Choffrut & Grigorieff 2006, Clemente et al. @ STACS 2017]

A First Idea

- Assume S is bounded at coordinate j, i.e., there is an $M \in \mathbb{N}$ such that $\vec{v}[j] \leq M$ for all $\vec{v} \in S$.



R and S can be partitioned according to their values in coordinate j.

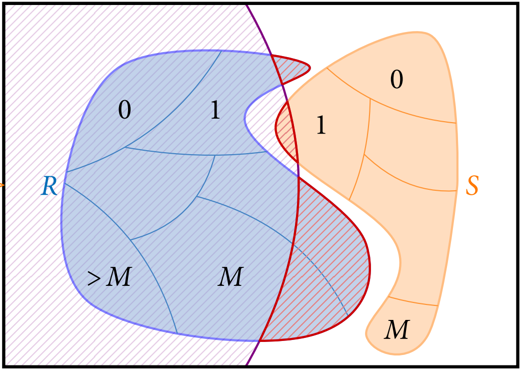
- Problem: Constructing R' and S' is expensive!

[Choffrut & Grigorieff 2006, Clemente et al. @ STACS 2017]

A First Idea

- Assume S is bounded at coordinate j, i.e., there is an $M \in \mathbb{N}$ such that $\vec{v}[j] \leq M$ for all $\vec{v} \in S$.

Assume that R and S are inseparable.

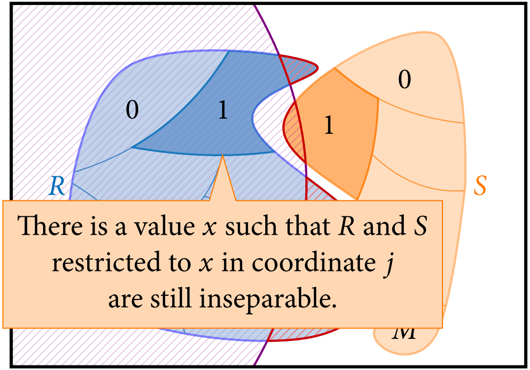


- Problem: Constructing R' and S' is expensive!

[Choffrut & Grigorieff 2006, Clemente et al. @ STACS 2017]

A First Idea

- Assume S is bounded at coordinate j, i.e., there is an $M \in \mathbb{N}$ such that $\vec{v}[j] \leq M$ for all $\vec{v} \in S$.

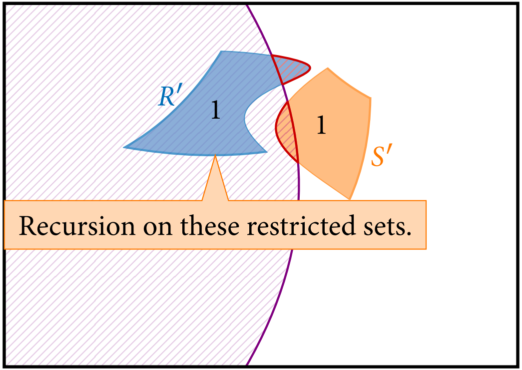


- Problem: Constructing R' and S' is expensive!

[Choffrut & Grigorieff 2006, Clemente et al. @ STACS 2017]

A First Idea

- Assume S is bounded at coordinate j, i.e., there is an $M \in \mathbb{N}$ such that $\vec{v}[j] \leq M$ for all $\vec{v} \in S$.



- Problem: Constructing R' and S' is expensive!

[Choffrut & Grigorieff 2006, Clemente et al. @ STACS 2017]

Twin-Unbounded Coordinates

- Let $R = U + P^*$ and $S = V + Q^*$.
- Repeat the following until stabilization. For each $1 \leq j \leq n$:
  - If S is bounded at j, remove all vectors $\vec{v} \in P$ with $\vec{v}[j] > 0$.
  - If R is bounded at j, remove all vectors $\vec{v} \in Q$ with $\vec{v}[j] > 0$.
- $\hat{P}$ and $\hat{Q}$ are the sets of all remaining vectors in P resp. Q after the procedure above.
- The remaining (unbounded) coordinates are called twin-unbounded.

Lemma

Let $R = U + P^*$ and $S = V + Q^*$ be two hyperlinear sets. Then the following are equivalent:

1. *R and S are inseparable by recognizable sets,*
2. $(R - \hat{P}^*) \cap (S - \hat{Q}^*) \neq \emptyset$, *and*
3. $(R + \hat{Q}^*) \cap (S + \hat{P}^*) \neq \emptyset$.

- Note that $R + \hat{Q}^*$ and $S + \hat{P}^*$ are still hyperlinear.
- Intersection non-emptiness of two semilinear sets is in NP. $\square$

Lemma

Let $R = U + P^*$ *and* $S = V + Q^*$ *be two hyperlinear sets. Then the following are equivalent:*

1. *R and S are inseparable by recognizable sets,*
2. $(R - \hat{P}^*) \cap (S - \hat{Q}^*) \neq \emptyset$, *and*
3. $(R + \hat{Q}^*) \cap (S + \hat{P}^*) \neq \emptyset$

R extended by the group generated by $\hat{P}$

- Note that $R + Q^*$ and $S + P^*$ are still hyperlinear.
- Intersection non-emptiness of two semilinear sets is in NP. □

Inseparability of Parikh Automata (1)

Main Theorem

1. The regular separability problem for Parikh automata is coNP-complete.
2. The recognizable separability problem for semilinear sets is coNP-complete.

- We show that regular inseparability of Parikh automata is in NP.
- Let $(\mathfrak{A}_1, C_1)$ and $(\mathfrak{A}_2, C_2)$ be two Parikh automata.
- Construct in polynomial time a DFA $\mathfrak{B}$ and semilinear sets D_1, D_2 such that $L(\mathfrak{A}_1, C_1)$ and $L(\mathfrak{A}_2, C_2)$ are regularly separable if, and only if, $L(\mathfrak{B}, D_1)$ and $L(\mathfrak{B}, D_2)$ are regularly separable.
- There are hyperlinear sets $R, S \subseteq \mathbb{N}^k$ such that $L(\mathfrak{B}, D_1)$ and $L(\mathfrak{B}, D_2)$ are regularly separable if, and only if, R and S are separable by a recognizable set.
  - In R and S we count the occurrences of (simple) cycles in accepting runs of $\mathfrak{B}$.
  - Dimension k is the number of all (simple) cycles in $\mathfrak{B}$.
  - Attention: This number can be exponential!

[Clemente et al. @ STACS 2017]

Inseparability of Parikh Automata (2)

- Recall $L(\mathfrak{B}, D_1)$ and $L(\mathfrak{B}, D_2)$ are regularly inseparable iff R and S are recognizably inseparable.
- We know: $R = U + P^*$ and $S = V + Q^*$ are inseparable iff $(R + \hat{Q}^*) \cap (S + \hat{P}^*) \neq \emptyset$
- Under-approximate $\hat{P}$ and $\hat{Q}$ by guessing a set of transitions participating in twin-unbounded coordinates.
  - Can be verified in NP.
- Construct in polynomial time Parikh automata $(\mathfrak{C}_1, E_1)$ and $(\mathfrak{C}_2, E_2)$ accepting sequences of cycles in $R + \hat{Q}^*$ resp. $S + \hat{P}^*$.
- Check whether the intersection of $L(\mathfrak{C}_1, E_1)$ and $L(\mathfrak{C}_2, E_2)$ is non-empty (in NP). □

Summary

Theorem

The following problems are coNP*-complete:*

1. *recognizable separability of semilinear sets.*
   - *also holds if the semilinear sets are given as existential or quantifier-free Presburger formulas*
2. *recognizable separability of two rational subsets of* $\Sigma^* \times \mathbb{N}^k$.
3. *regular separability of two Parikh-automata.*
4. *regularity of deterministic Parikh-automata with target sets given by quantifier-free Presburger formulas.*

Thank you!