# MIMUW Automata Theory Seminar 16/04/25

## ① Definition of the model

<u>Definition 1</u> A $\mathbb{Z}$-weighted automaton $\mathcal{A} = (Q, T, q_0, q_1)$ is an automaton where $Q$ is a finite set of states, $T \subseteq Q \times \mathbb{Z} \times Q$ is a finite set of transitions, and $q_0, q_1 \in Q$ are distinguished initial and final states, respectively.

We study three models:

| | Integer Semantics | Natural Semantics | VASS Semantics |
|---|---|---|---|
| swap: Configurations | $\mathbb{Z}$ | $\mathbb{N}$ | $\mathbb{Z}$ |
| Transitions | $\mathbb{Z}$ | $\mathbb{N}$ | $\mathbb{N}$ |

$p \xrightarrow{12} q \qquad (p, 5) \to (q, 17)$

## ② Motivating Question

Reachability: Input = ~~A $\mathbb{Z}$-weighted automaton initial configuration $(p, x)$, and a target configuration~~

(Binary encoding)

A $\mathbb{Z}$-weighted automaton, $p(x)$, $q(y)$

Question: $p(x) \xrightarrow{*} q(y)$ ?

<u>Theorem 2</u> Reachability in $\mathbb{Z}$-weighted automaton is NP-complete.

Coverability: Input = A $\mathbb{Z}$-weighted automaton, $p(x)$, $q(y)$

Question: $p(x) \xrightarrow{*} q(y')$ for some $y' \geq y$ ?

<u>Theorem 3</u> Coverability in $\mathbb{Z}$-weighted automaton is in $AC^1 \subseteq P$

<u>Question:</u> Why is reachability hard and coverability easy?

## ralised reachability problem

$\text{Reach}_{\mathbb{Z}}(S)$. Fixed: Semilinear set $S \subseteq \mathbb{Z}^{p+1}$ (one problem for each semilinear set).
Input: A $\mathbb{Z}$-weighted automaton $A = (Q, T, q_0, q_1)$, vector $\vec{t} \in \mathbb{Z}^p$
Question: Does there exist an $x \in \mathbb{Z}$ such that $q_0(0) \xrightarrow{*} q_1(x)$ and $(\vec{t}, x) \in S$.

---

④ **Examples.**

(1) $\cancel{S=\{t\}}$ $S = \{(t,x) : x = t\}$ (reachability) HARD

(2) $S = \{(t,x) : x \geq t\}$ (coverability) EASY

(3) $S = \{(t_1, t_2, x) : x \geq t_1 \text{ and } x \neq t_2\}$ (cover and avoid) EASY

(4) $S = \{x : x \equiv 1 \bmod 2 \text{ or } x = 0\}$ (odd or zero) HARD

(5) $S = \{(t,x) : x \in [t, 2t]\}$ (reach an interval) HARD

(6) $S = \{(t,x) : x \in (-\infty, 0] \cup [t, 2t]\}$ (reach a negative number or an interval) EASY

---

⑤ **Closer look at Example (6)**

Consider the following instance of "interval subset sum":

Input: $x_1, \ldots, x_n, t \in \mathbb{Z}$

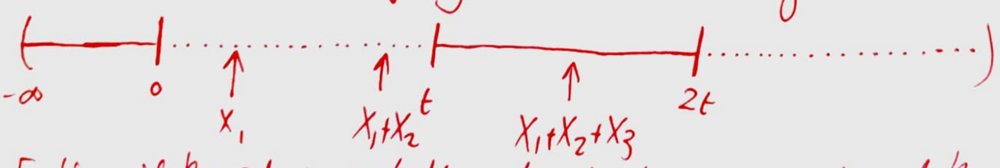
Question: $\exists I \subseteq [1,n] : \sum_{i \in I} x_i \in (-\infty, 0] \cup [t, 2t]$.

Polynomial-time algorithm: 1) If ~~and~~ $\exists i \in [1,n]: x_i \leq 0$ or $x_i \in [t, 2t]$, then the instance has the singleton solution $\{x_i\}$

2) Now, assuming all $x_i > 0$, one cannot use any item $x_i > 2t$.

3) Now, assuming all $x_i \in (0, t)$, compute $y = \sum_{i \in [1,n]} x_i$

If $y \geq t$, then return yes, else return no.




Feeling: if the set of reachable values is dense enough and the target is wide enough, then ~~even~~ it is easy to end up in the target set.

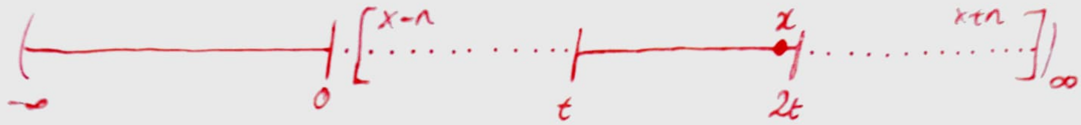
---

## y and main theorem

Definition 4 ~~(Density)~~ (Local density). Let $A \subseteq \mathbb{Z}$ be a semilinear set.

$$D_k(A, x) := \inf_{n \in \mathbb{N}} \frac{|A \cap (x + k[-n, n])|}{2n+1}$$

need $k$ for modulo classes




$$D(A) = \inf_{k \in \mathbb{N}} \inf_{x \in A} \left( D_k(A, x) \right)$$

$D((-\infty, 0] \cup [t, 2t]) \geq 1/4$

$A = (2\mathbb{Z}+1) \cup \{0\}$. $D_2(A, 0) = 0$ but $D_2(A, 1) = 1$

Now, for a semilinear set $S \subseteq \mathbb{Z}^{p+1}$, $D(S) = \inf_{\vec{z} \in \mathbb{Z}^p} \left( D(S[\vec{z}]) \right)$

$S[\vec{z}] = \{x \in \mathbb{Z} : (\vec{z}, x) \in S\}$.

## 7 Main Theorem

Theorem 5 Let ~~$\mathbb{Z}$~~ $S \subseteq \mathbb{Z}^{p+1}$ be semilinear then

(1) If $D(S) > 0$, the $\mathrm{Reach}_{\mathbb{Z}}(S)$ is in $AC^1$

(2) Otherwise, $\mathrm{Reach}_{\mathbb{Z}}(S)$ is in NP-complete.

we can now look back at the examples and use this dichotomy theorem to identify easy/hard problems.

## 8 Preliminaries to be mostly swept under the rug

(i) First, we reduce to the acyclic case.

Theorem 6. For each semilinear set $S \subseteq \mathbb{Z}^{p+d}$, the problem of $\mathrm{Reach}_{\mathbb{Z}}(S)$ reduces in logspace to $\mathrm{AcReach}_{\mathbb{Z}}(S)$. (Input $\mathbb{Z}$-weighted automaton is acyclic)

Proof idea: uses Eisenbrand and Shmonin's (2006) Caratheodory-like result for ILP / integer cones.

(ii) Remove modulo constraints from the given semilinear set $S$

Lemma 7. Let $S \subseteq \mathbb{Z}^{p+1}$ be a semilinear set. One can compute $B \geq 1$ so that for every $\vec{b} \in [0, B-1]^{p+1}$, $[S]_{B, \vec{b}} = \{\vec{v} \in \mathbb{Z}^{p+1} : B \cdot \vec{v} + \vec{b} \in S\}$ is effectively definable by a modulo-free Presburger formula. Moreover $D(S) > 0$ iff for all $\vec{b} \in [0, B-1]^{p+1}$, $D([S]_{B, \vec{b}}) > 0$.

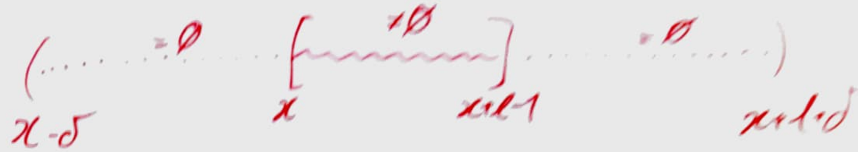
## Unbounded Isolation

**Definition 8.** Let $S \subseteq \mathbb{Z}^{p+1}$. We say $S$ has unbounded isolation if there exists an $\ell \geq 1$ such that for every $\delta \geq 1$, there exists $\vec{t} \in \mathbb{Z}^p$ and $x \in \mathbb{Z}$ such that

(a) $S[\vec{t}] \cap [x, x+\ell-1] \neq \emptyset$,

(b) $S[\vec{t}] \cap [x-\delta, x-1] = \emptyset$, and

(c) $S[\vec{t}] \cap [x+\ell, x+\ell+\delta] = \emptyset$.




**Lemma 9.** Suppose $S \subseteq \mathbb{Z}^{p+1}$ has unbounded isolation, then $\mathrm{Reach}_{\mathbb{Z}}(S)$ is NP hard.

Idea: You can use the value(s) in $[x, x+\ell-1]$ as target values in Subset Sum

---

## (10) Algorithmic Building Blocks.

We will define some "basic" sets for which $\mathrm{Reach}_{\mathbb{Z}}$ is in $AC^0$.

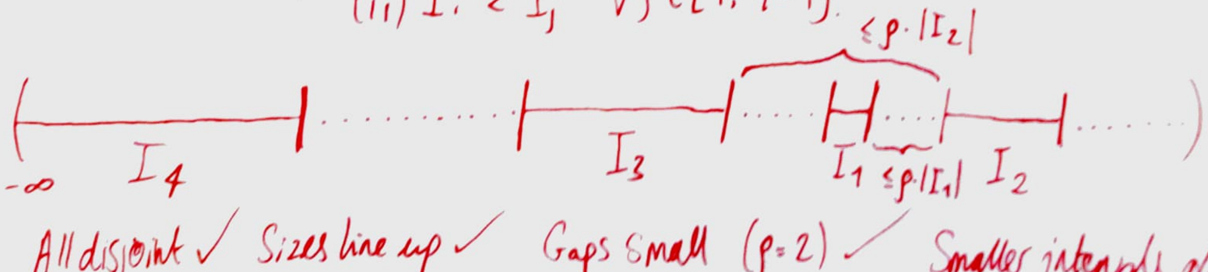
**Definition 10.** Let $\rho \geq 1$ be a rational constant. A sequence $I_1, \dots, I_{m+1}$ is called a $\rho$-chain of intervals if $I_{m+1}$ is one-sided infinite and

(A1) $I_1, \dots, I_m, I_{m+1}$ are pairwise disjoint,

(A2) $|I_1| \leq |I_2| \leq \dots \leq |I_m|$,

(A3) $\forall i \in [1,m]$, $d(I_i, I_{i+1}) \leq \rho \cdot |I_i|$, and

(A4) $\forall i \in [2, \cancel{m}\ m+1]$, (i) $I_i > I_j$ $\forall j \in [1, i-1]$, or
(ii) $I_i < I_j$ $\forall j \in [1, i-1]$.




All disjoint ✓ Sizes line up ✓ Gaps small ($\rho = 2$) ✓ Smaller intervals always on same side ✓

**Proposition 11.** Let $S \subseteq \mathbb{Z}^{p+1}$ (modulo-free semilinear). Then T.F.A.E.

(1) $D(S) > 0$

(2) $S$ does not have unbounded isolation

(3) $S$ is a finite union of ~~building blocks~~* (~~Basically ρ-chains~~) $\rho$-chains*

Proof ideas: (1) => (2) can be more easily seen as $\neg(2) \Rightarrow \neg(1)$. Point of isolation has 0 density.

(2) => (3) IS HARD WORK. (uses WQO theory for some strange matrices)

(3) => (1) Requires showing a $\rho$-chain has positive density.

## cient algorithm for $\rho$-chains

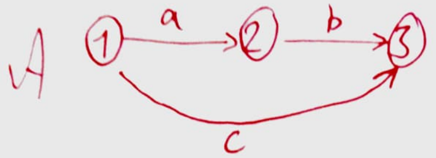
oposition 12 For every rational $\rho \geq 1$, and $m \geq 1$, $\mathrm{Reach}_{\mathbb{Z}}(\underbrace{S^{(\rho,m)}}_{\rho\text{ chain of } m+1 \text{ intervals}})$ is in $AC^1$

Run through proof ideas on our favourite example.

$$S = [-\infty, 0] \cup [t, 2t].$$

Step 1: transform problem into a matrix multiplication problem.

Consider $\mathcal{A}$




$a, b, c \in \mathbb{Z}$

$$\rightsquigarrow \begin{pmatrix} X^0 & X^a & X^c \\ 0 & X^0 & X^b \\ 0 & 0 & X^0 \end{pmatrix} =: A$$

$A$ is a matrix in $\mathrm{Mat}(|Q| \times |Q|, \underbrace{\mathbb{B}[X, X^{-1}]}_{\text{Laurent polynomials}})$

Polynomials in $X, X^{-1}$ over the Boolean semiring

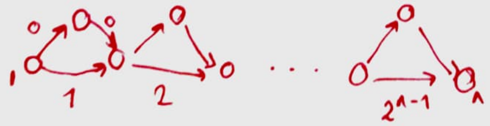
$1 \cdot 1 = 1$, $1 \cdot 0 = 0$
$1 + 1 = 1$, $1 + 0 = 1$.

Step 2: Recall $\mathcal{A}$ is acyclic so $A^n$ ($n = |Q|$) contains all the information we need for reachability.

$$(A^n)_{i,j} = \sum_{\substack{\text{paths } \pi \text{ from} \\ i \text{ to } j \text{ of length} \leq n}} X^{\mathrm{weight}(\pi)}.$$

Example: $(A^2)_{1,3} = X^0 \cdot X^c + X^a \cdot X^b + X^c \cdot X^0$
$= X^c + X^{a+b} + X^c$
$= X^c + X^{a+b}$

Problem: The polynomial $(A^n)_{i,j}$ may contain exponentially many terms




$$(A^n)_{1,n} = \sum_{e=0}^{2^n - 1} X^e$$

Solution for coverability: use maxplus semiring. (Alternately impose $X^i + X^j = X^{\max(i,j)}$)

Step 3: Impose some different equations that "preserve reachability".

(Eq. 1) $X^i + X^j + X^k \equiv X^{[i,j]} + X^k$ for all $j - i \in [0, k]$, and $k - j \in [0, k]$.

(Eq. 2) $X^h + X^i + X^j \equiv X^h + X^{[i,j]}$ for all $j - i \in [0, \infty)$ and $h - i \in [k, \infty)$

Definition: $X^{[i,j]} = X^i + X^{i+1} + \dots + X^j$.

4: Claim 13. Suppose $P \in \mathbb{B}[X, X^{-1}]$. Then there exists at most 2 (interval terms) $P_1, P_2$ such that $P \equiv_* P_1 + P_2$.

Proof: Assume $P$ contains at least 3 terms (o/w claim is trivially satisfied).
Let $l, r \in \mathbb{Z}$ be the least and greatest integers such that $P$ contains $X^l$ and $X^r$, respectively.

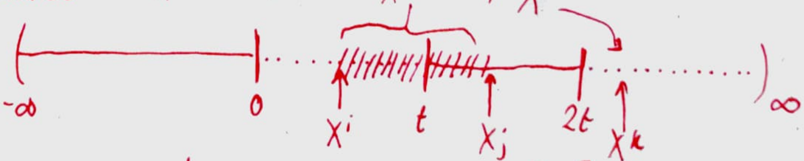
Case 1: $r - l \leq k$. Then $P \equiv_* X^{[l,r]}$ by (Eq. 1).

Case 2: $r - l > k$. Now, suppose there exists $m \in \mathbb{Z}$ such that $X^m$ is in $P$, $m - l \geq k$ and $l < m < r$.
Then, let $m$ be the least integer such that $X^m$ is in $P$, $m - l \geq k$, and $l < m < r$.
$\Rightarrow P \equiv_* X^l + P' + X^{[m,r]}$ by (Eq 2). where $P'$ is the polynomial containing the remaining terms.
Since $P'$ only contains terms $X^{m'}$ such that $m' - l < k$, $X^l + P' \equiv_* X^{[l,m']}$ by (Eq. 1)
where $X^{m'}$ is the highest order term in $P'$.
Thus $P \equiv_* X^{[l,m']} + X^{[m,r]}$

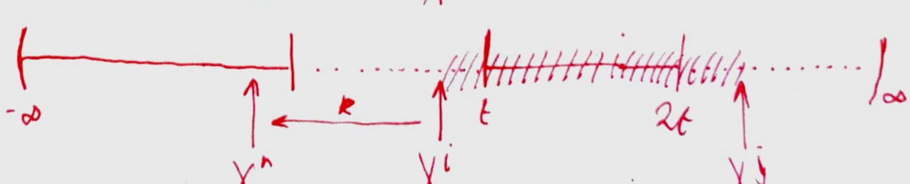
Now, suppose there does not exist $m \in \mathbb{Z}$ such that $X^m$ is in $P$, $m - l \geq k$, and $l < m < r$.
$\Rightarrow$ All terms $X^m$ in $P$ are "close" to $X^l$, i.e. $m - l < k$.
Let $m$ be the greatest integer such that $X^m$ is in $P$ and $l < m < r$.
Then, by (Eq. 1) $P \equiv_* X^{[l,m]} + X^r$. $\square$

Step 5: Claim 14. Let $P, P' \in \mathbb{B}[X, X^{-1}]$ such that $P \equiv_* P'$.
$P$ contains a term in $S = [-\infty, 0] \cup [t, 2t]$ iff $P'$ contains a term in $S$.

Proof Sketch: By pictures $X^i + X^j + X^k \equiv X^{[i,j]} + X^k$




$X^h + X^i + X^j \equiv X^h + X^{[i,j]}$



s: More generally:

Lemma 15. For a $\rho$-chain of $m+1$ intervals Every $P \in \mathbb{B}[X, X^{-1}]$ can be expressed as the sum of $(m\rho + 2m)^m$ (interval terms)

Step 7: Final comments: Given a matrix $A \in \mathrm{Mat}(n \times n, \mathbb{B}[X, X^{-1}])$ where every entry $(A)_{i,j}$ is written as the sum of $r \leq (m\rho + 2m)^m$ terms, we can compute $A^2$ in $AC^0$.

Then by using repeated matrix squaring $A, A^2, (A^2)^2, \dots, (A^2)^{\cdot^{\cdot^{2}}}\}^{\log(n)} = A^n$ one can decide whether $(A^n)_{1,n}$ contains a term in $S$ in $AC^1$.

## 12) Summary

- > Shown how to decide reachability into $(-\infty, 0] \cup [t, 2t]$ in $AC^1/P$.
- > How this roughly generalises to $\rho$-chains in general
- > I did not prove Proposition 11 (Split dense sets into $\rho$-chains)
- > Otherwise, if $D(S) = 0 \Rightarrow \mathrm{Reach}_{\mathbb{Z}}(S)$ is NP-hard.

## 13) Further results (Natural semantics)

Theorem 16. Let $S \subseteq \mathbb{Z}^\rho \times \mathbb{N}$ be a semilinear set

(1) if $D^+(S) > 0$, then $\mathrm{Reach}_{\mathbb{N}}(S)$ is in $AC^1$ $\rightarrow D^+_k(A, x) = \inf_{\substack{n \in \mathbb{N} \\ x - k \cdot n \in \mathbb{N}}} \frac{|A + (x + k[-n,n])|}{2n+1}$

(2) otherwise, $\mathrm{Reach}_{\mathbb{N}}(S)$ is NP-complete.

## 14) Further results (VASS semantics)

Definition 17. $A \subseteq \mathbb{N}$ is $\delta$-~~downwards~~ upwards closed if $A + \delta = \{a + \delta : a \in A\} \subseteq A$.

$A \subseteq \mathbb{N}$ is $(\delta, M)$-upwards closed if $\exists F \subseteq \mathbb{N}: |F| \leq M$ and $A \cup F$ is $\delta$-up. closed

$A \subseteq \mathbb{N}$ is quasi-upwards closed if $\exists \delta \geq 1, M \geq 0$: $A$ is $(\delta, M)$-up. closed

$S \subseteq \mathbb{Z}^\rho \times \mathbb{N}$ ~~$A \subseteq \mathbb{N}$~~ is uniformly quasi-upwards closed if $\forall \vec{t} \in \mathbb{Z}^\rho$, $\{x \in \mathbb{N} : (\vec{t}, x) \in S\}$ is quasi-upwards closed.

Theorem 18. Let $S \subseteq \mathbb{Z}^\rho \times \mathbb{N}$ be a semilinear set.

(1) if $S$ is uniformly quasi-upwards closed, $\mathrm{Reach}_{VASS}(S)$ is in $AC^1$.

(2) otherwise, $\mathrm{Reach}_{VASS}(S)$ is NP-complete.

## 14) Conclusion: All dichotomys are decidable.