Presburger Arithmetic
Reversal-Bounded Counter Machines

Stéphane Demri (demri@lsv.fr)

October 7th, 2016

Slides and lecture notes

http://www.lsv.fr/~demri/notes-de-cours.html

https://wikimpri.dptinfo.ens-cachan.fr/doku.php?id=cours:c-2-9-1

Plan of the lecture

- Previous lecture :
  - Introduction to Presburger arithmetic.
  - Decidability and quantifier elimination.
  - Automata-based approach.
- Presburger sets are the semilinear sets.
- Application: Parikh image of regular languages.
- Introduction to reversal-bounded counter machines.

The previous lecture in 2 slides (1/2)

- First-order theory $\mathrm{FO}(\mathbb{N})$ on $\langle \mathbb{N}, \leq, + \rangle$:

  $$\varphi ::= \top \mid \bot \mid t \leq t' \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists \mathsf{x}\ \varphi \mid \forall \mathsf{x}\ \varphi$$

- Presburger sets

  $$[\![\varphi(\mathsf{x}_1, \ldots, \mathsf{x}_n)]\!] \stackrel{\text{def}}{=} \{\langle \mathfrak{v}(\mathsf{x}_1), \ldots, \mathfrak{v}(\mathsf{x}_n)\rangle \in \mathbb{N}^n : \mathfrak{v} \models \varphi\}$$

- Quantifier-free fragment

  $$\top \mid \bot \mid t \leq t' \mid t \equiv_k t' \mid t = t' \mid t < t' \mid t \geq t' \mid t > t'$$

  (plus Boolean connectives)

- The satisfiability problem for the quantifier-free fragment is NP-complete.

Previous lecture in 2 slides (2/2)

- For every φ, there is a quantifier-free formula φ' such that
  1. $\mathit{free}(\varphi') \subseteq \mathit{free}(\varphi)$.
  2. φ' is logically equivalent to φ.
  3. φ' can be effectively built from φ.
- Presburger arithmetic is decidable.
- Alternative proof with the automata-based approach:

  "Presburger sets as regular languages of finite words"

Semilinear Sets

Formulae with one free variable

$$\varphi(\mathsf{x}) \stackrel{\text{def}}{=} (\mathsf{x} \neq 1 \wedge \mathsf{x} \neq 2) \wedge (\mathsf{x} = 0 \vee (\mathsf{x} \geq 3 \wedge \exists\, \mathsf{y}\ (\mathsf{x} = 3 + 2\mathsf{y})))$$

$$[\![\varphi(\mathsf{x})]\!] = \{0\} \cup \{3 + 2n : n \geq 0\}$$

- After the value 3, every two value belongs to $[\![\varphi(\mathsf{x})]\!]$.

$$\bullet \circ \circ \bullet \circ \bullet \circ \bullet \circ \bullet \circ \bullet \circ \bullet \circ \bullet \ldots$$

- This can be generalized.

$X \subseteq \mathbb{N}$ is ultimately periodic

$\stackrel{\text{def}}{\Leftrightarrow}$

there exist $N \geq 0$ and $P \geq 1$ such that for all $n \geq N$, we have $n \in X$ iff $n + P \in X$.

$$\overbrace{\bullet \circ \circ \bullet \circ \bullet \circ}^{\text{prefix}} \quad \overbrace{\bullet \bullet \circ \bullet \bullet}^{\text{period of length } P} \quad \bullet \bullet \circ \bullet \bullet \bullet \bullet \circ \bullet \bullet \bullet \bullet \circ \bullet \bullet \ldots$$

Examples of ultimately periodic sets

- The set of even numbers is ultimately periodic (with $N = 0$ and $P = 2$).
- The set of odd numbers is ultimately periodic (with $N = 0$ and $P = 2$).
- $[\![x \equiv_k k']\!]$ is ultimately periodic (with $N = 0$ and $P = k$).
- Ultimately periodic sets are closed under union, intersection and complementation.

Proof for complementation

- Suppose X is ultimately periodic and $\overline{X} = \mathbb{N} \setminus X$.
- The statements below are equivalent for $n \geq N$:
  - $n \in \overline{X}$,
  - $n \notin X$
    (by definition of $\overline{X}$),
  - $n + P \notin X$
    (X is ultimately periodic with parameters N and P),
  - $n + P \in \overline{X}$
    (by definition of $\overline{X}$).
- $\overline{X}$ is ultimately periodic too and the same parameters N and P can be used.

● ○ ○ ● ○ ● ○ ● ● ○ ● ● ● ● ○ ● ● ● ● ○ ● ● ● ● ○ ● ● . . .

○ ● ● ○ ● ○ ● ○ ○ ● ○ ○ ○ ○ ● ○ ○ ○ ○ ● ○ ○ ○ ○ ● ○ ○ . . .

Ultimately periodic sets X are Presburger sets

$$(\bigwedge_{k \in [0,N-1] \setminus X} x \neq k) \wedge [(\bigvee_{k \in [0,N-1] \cap X} x = k) \vee$$

$$((x \geq N) \wedge (\exists\, y \bigvee_{k \in [N,N+P-1] \cap X} (x = k + Py)))]$$

It remains to show the converse result.

Semilinear sets of dimension 1

For every formula $\varphi(\mathsf{x})$ with a unique free variable x, $[\![\varphi]\!]$ is an ultimately periodic set.

- Formula $\varphi(\mathsf{x})$ with a unique free variable x.
- φ': equivalent quantifier-free formula.
- φ' is a Boolean combination of atomic formulae of one of the forms below: $\top$, $\bot$, $\mathsf{x} \leq k$, $\mathsf{x} \equiv_k k'$.
- Each atomic formula defines an ultimately periodic set and ultimately periodic sets are closed under union, intersection and complementation.
- So $[\![\varphi']\!] = [\![\varphi]\!]$ is ultimately periodic.

Semilinear sets

- A linear set X is defined by a basis $\mathbf{b} \in \mathbb{N}^d$ and a finite set of periods $\mathfrak{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_m\} \subseteq \mathbb{N}^d$:

$$X = \{\mathbf{b} + \sum_{i=1}^{m} \lambda_i \mathbf{p}_i : \lambda_1, \ldots, \lambda_m \in \mathbb{N}\}$$

- A linear set:

$$\left\{ \begin{pmatrix} 3 \\ 4 \end{pmatrix} + i \times \begin{pmatrix} 2 \\ 5 \end{pmatrix} + j \times \begin{pmatrix} 4 \\ 7 \end{pmatrix} : i, j \in \mathbb{N} \right\}$$

- A semilinear set is a $\boxed{\text{finite}}$ union of linear sets.

- Each semilinear set can be represented by a finite set of pairs of the form $\langle \mathbf{b}, \mathfrak{P} \rangle$.

Ultimately periodic sets are semilinear sets

- Ultimately periodic set X with parameters N and P.

$$X = \Big(\bigcup_{n \in [0, N-1] \cap X} \{n\}\Big) \cup \Big(\bigcup_{n \in [N, N+P-1] \cap X} \{n + \lambda P : \lambda \in \mathbb{N}\}\Big)$$

- $\{n\}$ is a linear set with no period.
- $\{n + \lambda P : \lambda \in \mathbb{N}\}$ is a linear set with basis n and unique period P.

The fundamental characterisation

[Ginsburg & Spanier, PJM 66]

- For every Presburger formula φ with $d \geq 1$ free variables, $[\![\varphi]\!]$ is a semilinear subset of $\mathbb{N}^d$.
- For every semilinear set $X \subseteq \mathbb{N}^d$, there is φ such that $X = [\![\varphi]\!]$.
- The class of semilinear sets are effectively closed under union, intersection, complementation and projection.
- For instance, ($X_1 = [\![\varphi_1]\!]$ and $X_2 = [\![\varphi_2]\!]$) imply $X_1 \cap X_2 = [\![\varphi_1 \wedge \varphi_2]\!]$
- Presburger formula for

$$\left\{ \begin{pmatrix} 3 \\ 4 \end{pmatrix} + i \times \begin{pmatrix} 2 \\ 5 \end{pmatrix} + j \times \begin{pmatrix} 4 \\ 7 \end{pmatrix} : i, j \in \mathbb{N} \right\}$$

$$\exists \ \mathrm{y}, \mathrm{y}' \ (\mathrm{x}_1 = 3 + 2\mathrm{y} + 4\mathrm{y}' \wedge \mathrm{x}_2 = 4 + 5\mathrm{y} + 7\mathrm{y}')$$

$X = \{2^n : n \in \mathbb{N}\}$ is not a Presburger set

- *Ad absurdum*, suppose that X is semilinear.
- Since X is infinite, there are $\mathbf{b} \geq 0$ and $\mathbf{p}_1, \ldots, \mathbf{p}_m > 0$ ($m \geq 1$) such that

$$Y \stackrel{\text{def}}{=} \{\mathbf{b} + \sum_{i=1}^{m} \lambda_i \mathbf{p}_i : \lambda_1, \ldots, \lambda_m \in \mathbb{N}\} \subseteq X$$

- There exists $2^\alpha \in Y$ such that $\mathbf{p}_1 < 2^\alpha$.
- By definition of Y, we have $2^\alpha + \mathbf{p}_1 \in Y$.
- But, $2^\alpha < 2^\alpha + \mathbf{p}_1 < 2^{\alpha+1}$, contradiction.

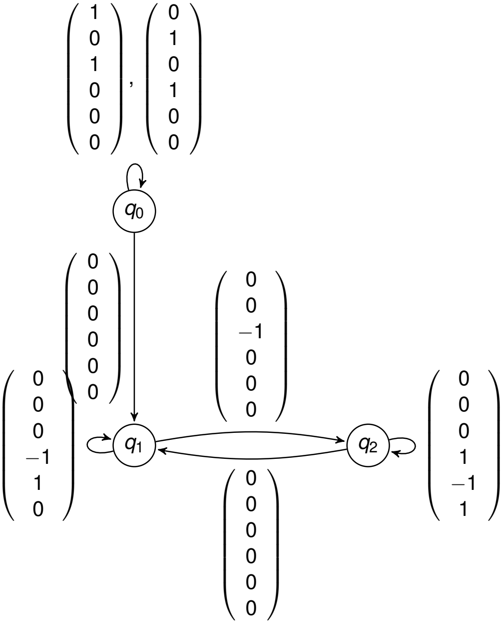
$X = \{n^2 : n \in \mathbb{N}\}$ is not a Presburger set

- *Ad absurdum*, suppose that X is semilinear.
- Since X is infinite, there are $\mathbf{b} \geq 0$ and $\mathbf{p}_1, \ldots, \mathbf{p}_m > 0$ ($m \geq 1$) such that

$$Z \stackrel{\text{def}}{=} \{\mathbf{b} + \sum_{i=1}^{m} \lambda_i \mathbf{p}_i : \lambda_1, \ldots, \lambda_m \in \mathbb{N}\} \subseteq X$$

- Let $N \in \mathbb{N}$ be such that $N^2 \in Z$ and $(2N + 1) > \mathbf{p}_1$.
- Since Z is a linear set, we also have $(N^2 + \mathbf{p}_1) \in Z$.
- However $(N + 1)^2 - N^2 = (2N + 1) > \mathbf{p}_1$.
- Hence $N^2 < N^2 + \mathbf{p}_1 < (N + 1)^2$, contradiction.

A VASS weakly computing multiplication



Weak multiplication

$$\left\{ \begin{pmatrix} a \\ b \\ f \end{pmatrix} \in \mathbb{N}^3 \;\middle|\; \exists \begin{pmatrix} c \\ d \\ e \end{pmatrix} \in \mathbb{N}^3,\ \langle q_0, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle \xrightarrow{*} \langle q_1, \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \rangle \right\} =$$

$$\left\{ \begin{pmatrix} n \\ m \\ p \end{pmatrix} \in \mathbb{N}^3 : p \leq n \times m \right\}.$$

Weak multiplication in a VASS

- Suppose there is $\varphi(\mathsf{x}_1, \ldots, \mathsf{x}_6)$ such that

$$[\![\varphi(\mathsf{x}_1, \ldots, \mathsf{x}_6)]\!] = \left\{ \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \;\middle|\; \langle q_0, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle \xrightarrow{*} \langle q_1, \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \rangle \right\}$$

- Formula $\psi(\mathsf{x})$ below verifies $[\![\psi(\mathsf{x})]\!] = \{n^2 \mid n \in \mathbb{N}\}$

$$\exists\, \mathsf{x}_1, \ldots, \mathsf{x}_5\ \varphi(\mathsf{x}_1, \ldots, \mathsf{x}_5, \mathsf{x}) \wedge \mathsf{x}_1 = \mathsf{x}_2 \wedge$$

$$\forall\, \mathsf{x}'\ (\mathsf{x}' > \mathsf{x}) \Rightarrow \neg\exists\, \mathsf{x}_3, \mathsf{x}_4, \mathsf{x}_5\ \varphi(\mathsf{x}_1, \ldots, \mathsf{x}_5, \mathsf{x}')$$

Contradiction!

Parikh Image of Regular Languages

Parikh image

- $\Sigma = \{a_1, \ldots, a_k\}$ with ordering $a_1 < \cdots < a_k$.
- Parikh image of $u \in \Sigma^*$: $\begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_k \end{pmatrix} \in \mathbb{N}^k$ where each n_j is the number of occurrences of a_j in u.
- Parikh image of $u = a\ b\ a\ a\ b$, written $\Pi(u)$, is $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$.
- Definition for Parikh image extends to languages.
- The Parikh image of any context-free language is semilinear. [Parikh, JACM 66]
- Effective computation from pushdown automata.

Bounded languages

- Language $\mathrm{L} \subseteq \Sigma^*$ bounded $\overset{\text{def}}{\Leftrightarrow}$

$$\mathrm{L} \subseteq u_1^* \cdots u_n^*$$

  for some words $u_1, \ldots, u_n$ in Σ^*.

- $\mathrm{L} \subseteq \Sigma^*$ is bounded and regular iff it is a finite union of languages of the form

$$u_0 v_1^* u_1 \cdots v_k^* u_k$$

- The Parikh images of bounded and regular languages are semilinear (i.e. Presburger sets).

Counting letters in bounded and regular languages

- Parikh image of $u_0 v_1^* u_1 \cdots v_k^* u_k$ is equal to

  $$\{\mathbf{b} + \lambda_1 \mathbf{p}_1 + \cdots \lambda_k \mathbf{p}_k : \lambda_1, \ldots, \lambda_k \in \mathbb{N}\}$$

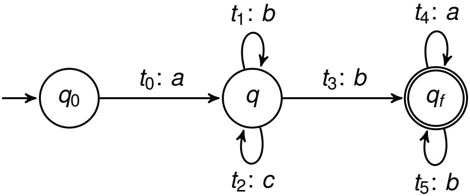
  with
  - $\mathbf{b} = \Pi(u_0) + \cdots + \Pi(u_k)$,
  - $\mathbf{p}_i = \Pi(v_i)$ for every $i \in [1, k]$.
- Finite union of such languages handled by finite unions of linear sets.
- Then, constructing a Presburger formula for the Parikh image easily follows.

Underapproximation by bounded languages

- For every regular language L, there is a bounded and regular language L' such that
  1. $\mathrm{L}' \subseteq \mathrm{L}$,
  2. $\Pi(\mathrm{L}') = \Pi(\mathrm{L})$.
- The proof consists in constructing L' effectively.
- $\mathcal{A} = \langle \Sigma, Q, Q_0, \delta, F \rangle$ such that $\mathrm{Lan}(\mathcal{A}) = \mathrm{L}$.

Paths, simple loops and extended paths

- Path π: finite sequence of transitions corresponding to a path in the control graph of $\mathcal{A}$.
- $\mathrm{first}(\pi)$ [resp. $\mathrm{last}(\pi)$]: first [resp. last] state of a path π.
- *lab*(π): label of π as a word of Σ^*.
- Simple loop *sl*: non-empty path that starts and ends by the same state and this is the only repeated state in it.
- "*sl* loops on its first state".
- Number of simple loops $\leq \mathrm{card}(\delta)^{\mathrm{card}(Q)}$.
- Arbitrary total linear ordering $\prec$ on simple loops.



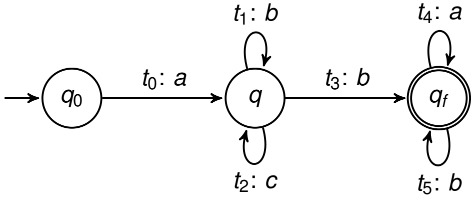
- Path $\pi = t_0 t_1 t_2 t_1 t_3$.
- Label $lab(\pi) = abcbb$.
- Simple loops $sl_1 = t_1$ and $sl_2 = t_2$.

Generalising the notion of path

- Encoding families of paths with extended paths.
- Extended path **P**:

$$\pi_0 \ S_1 \ \pi_1 \ \cdots \ S_\alpha \ \pi_\alpha$$

1. the S_i's are non-empty sets of simple loops,
2. the π_i's are non-empty paths,
3. if S occurs just before [resp. after] a path π, then all the simple loops in S loops on the first [resp. last] state of π.



$$t_0 \cdot t_1 \cdot \{t_1, t_2\} \cdot t_3 \cdot \{t_4, t_5\} \cdot t_4 \cdot t_5 \cdot t_5$$

Some more auxiliary notions

- Skeleton of **P** is the path $\pi_0 \cdots \pi_\alpha$.
- $S = \{sl_1, \ldots, sl_m\}$ with $sl_1 \prec \cdots \prec sl_m$

  $$e(S) \stackrel{\text{def}}{=} lab(sl_1)^+ \cdots lab(sl_m)^+$$

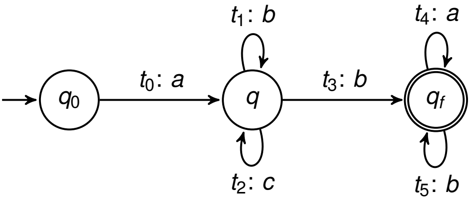
  (regular expression $e(S)$)
- $e(\mathbf{P}) \stackrel{\text{def}}{=} lab(\pi_0) \cdot e(S_1) \cdots e(S_\alpha) \cdot lab(\pi_\alpha)$.
- $\mathrm{Lan}(e)$: language defined by the regular expression e. $\mathrm{Lan}(e)$ is regular and bounded.
- $\mathrm{Lan}(\mathbf{P}) \stackrel{\text{def}}{=} \mathrm{Lan}(e(\mathbf{P}))$.
- When the first state occuring in the skeleton of **P** is in Q_0 and the last state is in F, then

  $$\mathrm{Lan}(e(\mathbf{P})) \subseteq \mathrm{Lan}(\mathcal{A})$$

Small extended path

- Small extended path:
  1. π_0 and π_α have at most $2 \times \mathrm{card}(Q)$ transitions,
  2. $\pi_1, \ldots, \pi_{\alpha-1}$ have at most $\mathrm{card}(Q)$ transitions,
  3. for each $q \in Q$, there is at most one set S containing simple loops on q.
- Length of the skeleton bounded by $\mathrm{card}(Q)(3 + \mathrm{card}(Q))$.
- The set of small extended paths is finite.

Example



- Small extended path **P**

$$t_0 \cdot t_1 \cdot \{t_1, t_2\} \cdot t_3 \cdot \{t_4, t_5\} \cdot t_4 \cdot t_5 \cdot t_5$$

- Regular expression $e(\mathbf{P})$ (with $t_1 \prec t_2$ and $t_5 \prec t_4$)

$$a \cdot b \cdot b^+ \cdot c^+ \cdot b \cdot b^+ \cdot a^+ \cdot a \cdot b \cdot b$$

How to proceed from a given run ρ

- Sequence of accepting extended paths $\mathbf{P}_0, \mathbf{P}_1, \ldots, \mathbf{P}_\beta$ such that
  - all the $\mathbf{P}_i$'s are accepting extended paths,
  - $\mathbf{P}_0$ is equal to ρ viewed as an extended path,
  - $\mathbf{P}_\beta$ is a small and accepting extended path,
  - $\mathbf{P}_{i+1}$ is obtained from $\mathbf{P}_i$ by removing a simple loop while $\Pi(\mathrm{Lan}(\mathbf{P}_i)) \subseteq \Pi(\mathrm{Lan}(\mathbf{P}_{i+1}))$.
- At the end of this process,

$$\Pi(\mathit{lab}(\rho)) \in \Pi(\mathrm{Lan}(\mathbf{P}_\beta)) \quad \text{and} \quad \Pi(\mathrm{Lan}(\mathbf{P}_\beta)) \subseteq \Pi(\mathrm{Lan}(\mathcal{A}))$$

From $\mathbf{P}_i$ to $\mathbf{P}_{i+1}$

$$\mathbf{P}_i = \pi_0 \; S_1 \; \pi_1 \; \cdots \; S_\alpha \; \pi_\alpha$$

(a) $\alpha \leq \mathrm{card}(Q)$,

(b) each path in $\pi_1, \ldots, \pi_{\alpha-1}$ have length less than $\mathrm{card}(Q)$,

(c) each state has at most one S_i with simple loops on it.

$\mathbf{P}_0$ verifies these conditions.

Three cases (1/2)

- $\mathbf{P}_i$ is a small extended path. We are done.

- $\pi_\alpha = \pi \cdot \textit{sl} \cdot \pi'$ where
  1. *sl* is a simple loop on q,
  2. $\pi\pi' \neq \varepsilon$,
  3. S_γ already contains simple loops on q.

  $\mathbf{P}_{i+1}$ is equal to:

$$\pi_0 \ \cdots \ S_{\gamma-1} \ \pi_{\gamma-1} \ (S_\gamma \cup \{\textit{sl}\}) \ \cdots \ \pi_{\alpha-1} \ S_\alpha \ (\pi\pi')$$

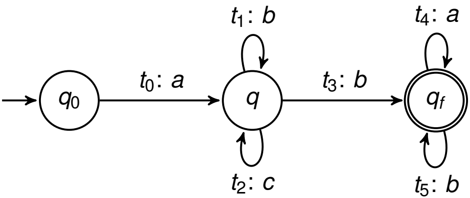
Three cases (2/2)

- $\pi_\alpha = \pi \cdot \mathit{sl} \cdot \pi'$ where
  1. sl is a simple loop on q,
  2. the first one occurring in $\pi \cdot \mathit{sl}$,
  3. $\pi\pi' \neq \varepsilon$,
  4. no S_γ already contains simple loops on q.

  $\mathbf{P}_{i+1}$ is equal to: $\pi_0 \ \cdots \ S_\alpha \ \pi \ \{\mathit{sl}\} \ \pi'$.

- Three properties easy to prove:
  1. $\Pi(\mathrm{Lan}(\mathbf{P}_i)) \subseteq \Pi(\mathrm{Lan}(\mathbf{P}_{i+1}))$.
  2. $\mathbf{P}_{i+1}$ satisfies the three previous conditions.
  3. $\mathrm{Lan}(\mathbf{P}_{i+1}) \subseteq \mathrm{Lan}(\mathcal{A})$.

Example



$$t_0 \cdot (t_1)^7 \cdot (t_2)^7 (t_1)^8 \cdot t_3 \cdot (t_4)^7 \cdot (t_5)^7 \cdot (t_4)^8$$

- $\mathbf{P}_{22} = t_0 \cdot \{t_1, t_2\} \cdot t_3 \cdot (t_4)^7 \cdot (t_5)^7 \cdot (t_4)^8$.
- $\mathbf{P}_{38} = t_0 \cdot \{t_1, t_2\} \cdot t_3 \cdot \{t_4, t_5\} \cdot (t_4)^6$.
- $\mathbf{P}_{38}$ is a small extended path.

Time to conclude!

- FSA $\mathcal{A}$ over a k-size alphabet Σ. One can compute a formula $\varphi_{\mathcal{A}}(\mathsf{x}_1, \ldots, \mathsf{x}_k)$ in $\mathrm{FO}(\mathbb{N})$ such that

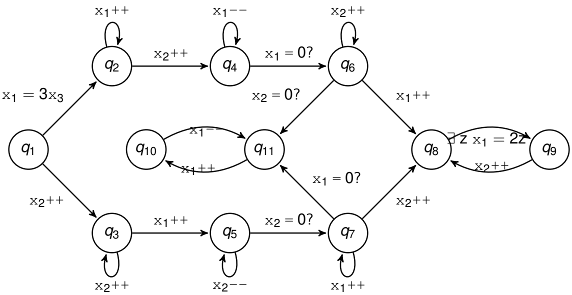
$$\Pi(\mathrm{Lan}(\mathcal{A})) = [\![\varphi_{\mathcal{A}}]\!]$$

- $\mathrm{Lan}(\mathcal{A})$ includes a bounded and regular language L with the same Parikh image.
- L can be computed by enumerating the regular expressions obtained from small and accepting extended paths and then check inclusion with $\mathrm{Lan}(\mathcal{A})$.
- Disjunction made of the formulae obtained for each bounded and regular language included in $\mathrm{Lan}(\mathcal{A})$.

Presburger Counter Machines

Presburger counter machines (PCM)

- Presburger counter machine $\mathcal{M} = \langle Q, T, C \rangle$:
  - Q is a nonempty finite set of control states.
  - C is a finite set of counters $\{\mathrm{x}_1, \ldots, \mathrm{x}_d\}$ for some $d \geq 1$.
  - T = finite set of transitions of the form $t = \langle q, \varphi, q' \rangle$ where $q, q' \in Q$ and φ is a Presburger formula with free variables $\mathrm{x}_1, \ldots, \mathrm{x}_d, \mathrm{x}'_1, \ldots, \mathrm{x}'_d$.

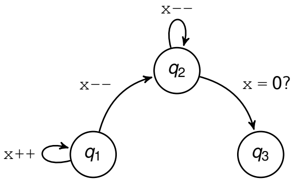


- Configuration $\langle q, \mathbf{x} \rangle \in Q \times \mathbb{N}^d$.

Transition system $\mathfrak{T}(\mathcal{M})$

- Transition system $\mathfrak{T}(\mathcal{M}) = \langle Q \times \mathbb{N}^d, \rightarrow \rangle$:

$$\langle q, \mathbf{x} \rangle \rightarrow \langle q', \mathbf{x}' \rangle \quad \stackrel{\text{def}}{\Leftrightarrow} \quad \text{there is } t = \langle q, \varphi, q' \rangle \text{ s.t. } \mathfrak{v}[\overline{\mathtt{x}} \leftarrow \mathbf{x}, \overline{\mathtt{x}'} \leftarrow \mathbf{x}'] \models \varphi$$



$$\begin{tikzcd}
\langle q_1, 0 \rangle \arrow[r] & \langle q_1, 1 \rangle \arrow[r] \arrow[d] & \langle q_1, 2 \rangle \arrow[r] \arrow[d] & \langle q_1, 3 \rangle \arrow[r] \arrow[d] & \langle q_1, 4 \rangle \arrow[d] & \cdots \\
 & \langle q_2, 0 \rangle \arrow[d] & \langle q_2, 1 \rangle \arrow[l] & \langle q_2, 2 \rangle \arrow[l] & \langle q_2, 3 \rangle \arrow[l] & \cdots \\
 & \langle q_3, 0 \rangle & & & &
\end{tikzcd}$$

- $\xrightarrow{*}$: reflexive and transitive closure of $\rightarrow$.

Decision problems

- Reachability problem:

  Input: PCM $\mathcal{M}$, $\langle q_0, \mathbf{x}_0 \rangle$ and $\langle q_f, \mathbf{x}_f \rangle$.

  Question: $\langle q_0, \mathbf{x}_0 \rangle \xrightarrow{*} \langle q_f, \mathbf{x}_f \rangle$?

- Control state reachability problem:

  Input: PCM $\mathcal{M}$, $\langle q_0, \mathbf{x}_0 \rangle$ and q_f.

  Question: $\exists \mathbf{x}_f \; \langle q_0, \mathbf{x}_0 \rangle \xrightarrow{*} \langle q_f, \mathbf{x}_f \rangle$?

- Control state repeated reachability problem:

  Input: PCM $\mathcal{M}$, $\langle q_0, \mathbf{x}_0 \rangle$ and q_f.

  Question: is there an infinite run starting from $\langle q_0, \mathbf{x}_0 \rangle$ such that the control state q_f is repeated infinitely often?

- Boundedness problem:

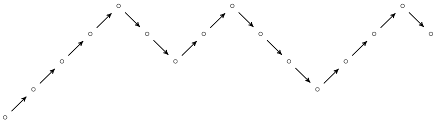
  Input: PCM $\mathcal{M}$ and $\langle q_0, \mathbf{x}_0 \rangle$.

  Question: is the set of configurations reachable from $\langle q_0, \mathbf{x}_0 \rangle$ finite?

What is Reversal-Boundedness?

Reversal-bounded counter machines

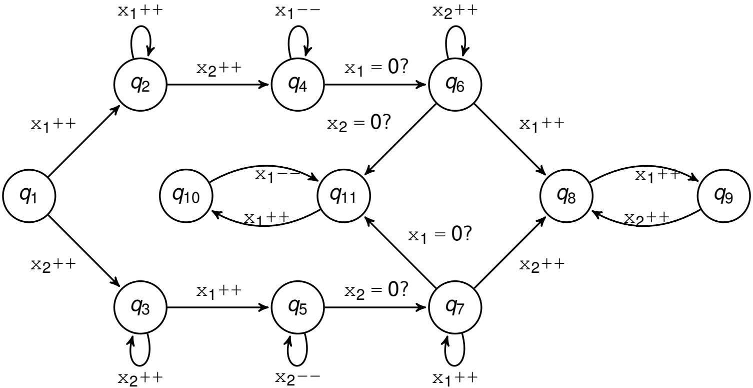
- Reversal: Alternation from a nonincreasing mode to a nondecreasing mode and vice-versa.



- Sequence with 3 reversals:

$$0011223334444\overline{3}33222\overline{3}33444455555 5\overline{4}$$

- A run is r-reversal-bounded whenever the number of reversals of each counter is less or equal to r.



$$\varphi = (x_1 \geq 2 \wedge x_2 \geq 1 \wedge (x_2 + 1 \geq x_1) \vee (x_2 \geq 2 \wedge x_1 \geq 1 \wedge x_1 + 1 \geq x_2)$$

$$[\![\varphi]\!] = \{\mathbf{y} \in \mathbb{N}^2 : \langle q_1, \mathbf{0} \rangle \xrightarrow{*} \langle q_9, \mathbf{y} \rangle\}$$

Presburger-definable reachability sets

- Let $\langle \mathcal{M}, \langle q_0, \mathbf{x}_0 \rangle \rangle$ be r-reversal-bounded for some $r \geq 0$. For each control state q, the set

$$R = \{\mathbf{y} \in \mathbb{N}^d : \exists \text{ run } \langle q_0, \mathbf{x}_0 \rangle \xrightarrow{*} \langle q, \mathbf{y} \rangle\}$$

  is effectively semilinear [Ibarra, JACM 78].

- One can compute effectively a Presburger formula φ such that $[\![\varphi]\!] = R$.

- The reachability problem with bounded number of reversals:

  Input: PCM $\mathcal{M}$, $\langle q, \mathbf{x} \rangle$, $\langle q', \mathbf{x}' \rangle$ and $r \geq 0$.

  Question: Is there a run $\langle q, \mathbf{x} \rangle \xrightarrow{*} \langle q', \mathbf{x}' \rangle$ s.t. each counter performs during the run a number of reversals bounded by r?

- The problem is decidable for a large class of counter machines.

Features of the proof

- Reachability relation of simple loops can be expressed in Presburger arithmetic.
- Runs can be normalized so that:
  - each simple loop is visited at most a doubly-exponential number of times,
  - the different simple loops are visited in a structured way.

Current class of counter machines $\mathcal{M} = \langle Q, T, C \rangle$

- Q is a finite set of control states and $C = \{\mathrm{x}_1, \ldots, \mathrm{x}_d\}$.
- T is a finite set of transitions.
- Each transition is labelled by $\langle g, \mathbf{a} \rangle$ where $\mathbf{a} \in \mathbb{Z}^d$ (update) and g is a guard following

  $$g ::= \top \mid \perp \mid \mathrm{x} \sim k \mid g \wedge g \mid g \vee g \mid \neg g$$

  where $\mathrm{x} \in C$, $\sim \in \{\leq, \geq, =\}$ and $k \in \mathbb{N}$.
- Update functions are those for VASS.
- Guards are more general than those for Minsky machines.
- Minsky machines and VASS belong to this class.

Mode vectors
– counter values for reversals –

- From a run

$$\rho = \langle q_0, \mathbf{x}_0 \rangle \xrightarrow{t_1} \langle q_1, \mathbf{x}_1 \rangle, \ldots$$

  we define mode vectors $\mathfrak{md}_0, \mathfrak{md}_1, \ldots$ such that each $\mathfrak{md}_i \in \{\mathrm{INC}, \mathrm{DEC}\}^d$.

- By convention, $\mathfrak{md}_0$ is the unique vector in $\{\mathrm{INC}\}^d$.

- For all $j \geq 0$ and for all $i \in [1, d]$, we have
  1. $\mathfrak{md}_{j+1}(i) \stackrel{\text{def}}{=} \mathfrak{md}_j(i)$ when $\mathbf{x}_j(i) = \mathbf{x}_{j+1}(i)$.
  2. $\mathfrak{md}_{j+1}(i) \stackrel{\text{def}}{=} \mathrm{INC}$ when $\mathbf{x}_{j+1}(i) - \mathbf{x}_j(i) > 0$.
  3. $\mathfrak{md}_{j+1}(i) \stackrel{\text{def}}{=} \mathrm{DEC}$ when $\mathbf{x}_{j+1}(i) - \mathbf{x}_j(i) < 0$.

- Number of reversals:

$$Rev_i \stackrel{\text{def}}{=} \{j \in [0, |\rho| - 1] : \mathfrak{md}_j(i) \neq \mathfrak{md}_{j+1}(i)\}$$

Reversal-boundedness formally

- Run ρ is r-reversal-bounded with respect to $i \overset{\text{def}}{\Leftrightarrow}$ $\mathrm{card}(Rev_i) \leq r$.
- Run ρ is r-reversal-bounded $\overset{\text{def}}{\Leftrightarrow}$ for every $i \in [1, d]$, we have $\mathrm{card}(Rev_i) \leq r$.
- $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$ is r-reversal-bounded $\overset{\text{def}}{\Leftrightarrow}$ every run from $\langle q, \mathbf{x} \rangle$ is r-reversal-bounded.
- $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$ is reversal-bounded $\overset{\text{def}}{\Leftrightarrow}$ there is some $r \geq 0$ such that every run from $\langle q, \mathbf{x} \rangle$ is r-reversal-bounded.

Semantical restriction

- $\mathcal{M}$ is uniformly reversal-bounded $\overset{\text{def}}{\Leftrightarrow}$ there is $r \geq 0$ such that for every initial configuration, the initialized counter machine is r-reversal-bounded.
- In the sequel, reversal-bounded counter machines come with a maximal number of reversals $r \geq 0$.
- Reversal-boundedness is essentially a semantical restriction on the runs.
- Reversal-boundedness detection problem on VASS is EXPSPACE-complete (the bound r can be computed).
- Reversal-boundedness detection problem on Minsky machines is undecidable.

Structure of the forthcoming proof

- Design a notion of extended path for which no reversal occurs and satisfaction of the guards remains constant.
- Any finite r-reversal-bounded run can be generated by a small sequence of such small extended paths.
- Reachability relation generated by any extended path is definable in Presburger arithmetic.

Intervals

- $\mathcal{M} = \langle Q, T, C \rangle$ with negation-free guards.
- *AG*: set of atomic guards of the form $\mathtt{x} \sim k$ occurring in $\mathcal{M}$.
- $\mathcal{K} = \{0 = k_1 < k_2 < \cdots < k_K\}$ and $K = \mathrm{card}(\mathcal{K})$.
- $\mathcal{I}$: set of non-empty intervals

$$\{[k_1, k_1], [k_1 + 1, k_2 - 1], [k_2, k_2], [k_2 + 1, k_3 - 1], [k_3, k_3], \ldots,$$
$$[k_K, k_K], [k_K + 1, +\infty)\} \setminus \{\emptyset\}$$

- At most $2K$ intervals and at least $K + 1$ intervals.

Counter values symbolically

- Linear ordering on $\mathcal{I}$ (for non-empty intervals):

$$[k_1, k_1] \leq [k_1{+}1, k_2{-}1] \leq [k_2, k_2] \leq [k_2{+}1, k_3{-}1] \leq [k_2, k_2] \leq \ldots$$

$$\ldots \leq [k_K, k_K] \leq [k_K + 1, +\infty)\}$$

- Interval map $\mathfrak{im} : C \rightarrow \mathcal{I}$.

- Distinct values from the same interval satisfy the same guards.

- Symbolic satisfaction relation $\mathfrak{im} \vdash g$:
  - $\mathfrak{im} \vdash g_1 \vee g_2 \overset{\text{def}}{\Leftrightarrow} \mathfrak{im} \vdash g_1$ or $\mathfrak{im} \vdash g_2$.
  - $\mathfrak{im} \vdash g_1 \wedge g_2 \overset{\text{def}}{\Leftrightarrow} \mathfrak{im} \vdash g_1$ and $\mathfrak{im} \vdash g_2$.
  - $\mathfrak{im} \vdash \mathrm{x} = k \overset{\text{def}}{\Leftrightarrow} \mathfrak{im}(\mathrm{x}) = [k, k]$.
  - $\mathfrak{im} \vdash \mathrm{x} \geq k \overset{\text{def}}{\Leftrightarrow} \mathfrak{im}(\mathrm{x}) \subseteq [k, +\infty)$.
  - $\mathfrak{im} \vdash \mathrm{x} \leq k \overset{\text{def}}{\Leftrightarrow} \mathfrak{im}(\mathrm{x}) \subseteq [0, k]$.

Completeness

- Interval maps and guards are built over the same set of constants.
- $\mathfrak{im} \vdash g$ can be checked in polynomial time in the sum of the respective sizes of $\mathfrak{im}$ and g.
- $\mathfrak{im} \vdash g$ iff for all $\mathfrak{f} : C \to \mathbb{N}$ and for all $\mathrm{x} \in C$, we have $\mathfrak{f}(\mathrm{x}) \in \mathfrak{im}(\mathrm{x})$ implies $\mathfrak{f} \models g$ (in Presburger arithmetic).

Guarded modes

- Guarded mode $\mathfrak{gmd}$ is a pair $\langle \mathfrak{im}, \mathfrak{md} \rangle$ where
  - $\mathfrak{im}$ is an interval map,
  - $\mathfrak{md} \in \{\mathrm{INC}, \mathrm{DEC}\}^d$.
- $t = q \xrightarrow{\langle g, \mathbf{a} \rangle} q'$ is compatible with $\mathfrak{gmd}$ $\overset{\text{def}}{\Leftrightarrow}$
  1. $\mathfrak{im} \vdash g$,
  2. for every $i \in [1, d]$,
     - $\mathfrak{md}(i) = \mathrm{INC}$ implies $\mathbf{a}(i) \geq 0$,
     - $\mathfrak{md}(i) = \mathrm{DEC}$ implies $\mathbf{a}(i) \leq 0$.

“*Bis repetita placent*”

- Path π is a sequence of transitions

$$q_1 \xrightarrow{\langle g_1, \mathbf{a}_1 \rangle} q'_1, \ldots, q_n \xrightarrow{\langle g_n, \mathbf{a}_n \rangle} q'_n$$

so that for every $i \in [1, n]$, we have $q'_i = q_{i+1}$.

- The effect of π is the update $\mathfrak{ef}(\pi) \stackrel{\text{def}}{=} \sum_j \mathbf{a}_j \in \mathbb{Z}^d$.
- Simple loop *sl* is a non-empty path that starts and ends by the same state and that's the only repeated state.
- Number of simple loops is $\leq \mathrm{card}(T)^{\mathrm{card}(Q)}$.
- Arbitrary total linear ordering $\prec$ on simple loops.

Extended path (bis)

- Extended path **P**:

$$\pi_0 \; S_1 \; \pi_1 \; \cdots \; S_\alpha \; \pi_\alpha$$

1. the S_i's are non-empty sets of simple loops,
2. the π_i's are non-empty paths,
3. if S occurs just before [resp. after] a path π, then all the simple loops in S loops on the first [resp. last] state of π.

Some more auxiliary notions

- A sequence of transitions is compatible with the guarded mode $\mathfrak{gmd} \overset{\text{def}}{\Leftrightarrow}$ all its transitions are compatible with $\mathfrak{gmd}$.
- Skeleton of **P** is the path $\pi_0 \cdots \pi_\alpha$.
- $S = \{sl_1, \ldots, sl_m\}$ with $sl_1 \prec \cdots \prec sl_m$

$$e(S) \stackrel{\text{def}}{=} (sl_1)^+ \cdots (sl_m)^+$$

  (the underlying alphabet is T)
- $e(\mathbf{P}) \stackrel{\text{def}}{=} \pi_0 \cdot e(S_1) \cdots e(S_\alpha) \cdot \pi_\alpha$.
- $\mathrm{Lan}(\mathbf{P}) \stackrel{\text{def}}{=} \mathrm{Lan}(e(\mathbf{P}))$.
- Run $\rho = \langle q_0, \mathbf{x}_0 \rangle \xrightarrow{t_1} \cdots \xrightarrow{t_\ell} \langle q_\ell, \mathbf{x}_\ell \rangle$ respects **P** $\overset{\text{def}}{\Leftrightarrow}$ $\pi = t_1 \cdots t_\ell \in \mathrm{Lan}(\mathbf{P})$.

Global phases (Intervals may change)

- Global phase: finite sequence of transitions such that each transition in it is compatible with some guarded mode $\langle \mathfrak{im}, \mathfrak{md} \rangle$, for some mode $\mathfrak{md} \in \{\mathrm{INC}, \mathrm{DEC}\}^d$.

- A run respecting a global phase has no reversal for all the counters (i.e. constant vector mode).

- r-reversal-bounded run $\rho = \langle q_0, \mathbf{x}_0 \rangle \cdots \langle q_\ell, \mathbf{x}_\ell \rangle$.
  - ρ can be divided as a sequence of subruns $\rho = \rho_1 \cdot \rho_2 \cdots \rho_L$.
  - Each ρ_i respects a global phase.
  - $L \leq (d \times r) + 1$.

Local phases

- Local phase: finite sequence of transitions such that each transition in it is compatible with some guarded mode $\langle \mathfrak{im}, \mathfrak{md} \rangle$.
- A run respecting a local phase has no reversals and the counter values satisfy the same atomic guards.
- r-reversal-bounded run $\rho = \langle q_0, \mathbf{x}_0 \rangle \cdots \langle q_\ell, \mathbf{x}_\ell \rangle$.
  - ρ can be divided as a sequence $\rho = \rho_1 \cdot \rho_2 \cdots \rho_{L'}$.
  - Each ρ_i respects a local phase.
  - $L' \leq ((d \times r) + 1) \times 2Kd$.

Sequences of extended paths

- $\mathbf{P}_1 \cdots \mathbf{P}_{L'}$ such that
  - each $\mathbf{P}_i$ is an extended path compatible with some guarded mode,
  - $\mathbf{P}_1 \cdots \mathbf{P}_{L'}$ is compatible with the control graph of $\mathcal{M}$.
- Any r-reversal-bounded run $\rho = \langle q_0, \mathbf{x}_0 \rangle \cdots \langle q_\ell, \mathbf{x}_\ell \rangle$ respects a sequence of extended paths $\mathbf{P}_1 \cdots \mathbf{P}_{L'}$ with

$$L' \leq ((d \times r) + 1) \times 2Kd$$

Small extended path (bis)

- Small extended path:
  1. π_0 and π_α have at most $2 \times \mathrm{card}(Q)$ transitions,
  2. $\pi_1, \ldots, \pi_{\alpha-1}$ have at most $\mathrm{card}(Q)$ transitions,
  3. for each $q \in Q$, there is at most one set S containing simple loops on q.
- Length of the skeleton bounded by $\mathrm{card}(Q)(3 + \mathrm{card}(Q))$.
- The set of small extended paths is finite.

Runs in normal form

- Run $\rho = \langle q_0, \mathbf{x}_0 \rangle \cdots \langle q_\ell, \mathbf{x}_\ell \rangle$ respecting **P** compatible with some guarded mode $\mathfrak{gmd}$.
- Then, there is small **P′** still compatible with $\mathfrak{gmd}$ and a run

$$\rho' = \langle q_0, \mathbf{x}_0 \rangle \cdots \langle q_\ell, \mathbf{x}_\ell \rangle$$

  such that ρ' respects **P′**.
- Generalization of the case for finite-state automata but with constraints on initial and final counter values.
- Convexity of the guards is used.

Small extended path compatible with $\mathfrak{gmd}$

- Extended path **P**:

$$\pi_0 \; S_1 \; \pi_1 \; \cdots \; S_\alpha \; \pi_\alpha$$

- Small extended path:
  1. π_0 and π_α have at most $2 \times \mathrm{card}(Q)$ transitions,
  2. $\pi_1, \ldots, \pi_{\alpha-1}$ have at most $\mathrm{card}(Q)$ transitions,
  3. for each $q \in Q$, there is at most one set S containing simple loops on q.

- For every transition $t = q \xrightarrow{\langle g, \mathbf{a} \rangle} q'$:
  1. $\mathfrak{im} \vdash g$,
  2. for every $i \in [1, d]$,
     - $\mathfrak{md}(i) = \mathrm{INC}$ implies $\mathbf{a}(i) \geq 0$,
     - $\mathfrak{md}(i) = \mathrm{DEC}$ implies $\mathbf{a}(i) \leq 0$.

Normal forms

- r-reversal-bounded run $\rho = \langle q_0, \mathbf{x}_0 \rangle \cdots \langle q_\ell, \mathbf{x}_\ell \rangle$.
- ρ can be divided as a sequence $\rho = \rho_1 \cdot \rho_2 \cdots \rho_{L'}$ such that
  - each ρ_i respects a small extended path $\mathbf{P}_i$ compatible with some guarded mode $\mathfrak{gmd}_i$.
  - $L' \leq ((d \times r) + 1) \times 2Kd$.

Reachability Sets are Presburger Sets

- Small extended path $\mathbf{P}$ compatible with $\mathfrak{gmd} = \langle \mathfrak{im}, \mathfrak{md} \rangle$

$$\pi_0 \ \{sl_1^1, \ldots, sl_1^{n_1}\} \ \pi_1 \ \cdots \ \{sl_\alpha^1, \ldots, sl_\alpha^{n_\alpha}\} \ \pi_\alpha$$

  where q_0 is the first control state in π_0 and q_f is the last control state in π_α $(= \pi'_\alpha \cdot t)$.

- There is $\varphi(\bar{\mathrm{x}}, \bar{\mathrm{y}})$ of exponential size in $|\mathcal{M}|$ such that

$$[\![\varphi]\!] = \{\langle \mathbf{x}_0, \mathbf{y} \rangle : \text{ there is a run } \langle q_0, \mathbf{x}_0 \rangle \xrightarrow{*} \langle q_f, \mathbf{y} \rangle \text{ respecting } \mathbf{P}\}$$

- φ states the following properties:
  1. the values in $\mathbf{x}_0$ belong to the right intervals induced by $\mathfrak{im}$,
  2. the counter values for the penultimate configuration $\langle q'_f, \mathbf{y}' \rangle$ belong to the right intervals induced by $\mathfrak{im}$,
  3. the values for $\bar{\mathrm{y}}$ are obtained from $\bar{\mathrm{x}}$ by considering the effects of the paths π_i plus a finite amount of times the effects of each simple loop occurring in $\mathbf{P}$.

Arghhhh !!!!!

$$\exists\, z_1^1,\ldots,z_1^{n_1},\ldots,z_\alpha^1,\ldots,z_\alpha^{n_\alpha}$$

$$(z_1^1 \geq 1) \wedge \cdots \wedge (z_1^{n_1} \geq 1) \wedge \cdots \wedge (z_\alpha^1 \geq 1) \wedge \cdots \wedge (z_\alpha^{n_\alpha} \geq 1) \wedge$$

$$(\bar{y} = \bar{x} + \mathfrak{ef}(\pi_0) + \cdots + \mathfrak{ef}(\pi_\alpha) + \sum_{i,j} z_i^j \mathfrak{ef}(sl_i^j)) \wedge$$

$$(\bigwedge_{\mathfrak{im} \vdash \mathrm{x}_c \sim k} x_c \sim k) \wedge (\bigwedge_{\text{not } \mathfrak{im} \vdash \mathrm{x}_c \sim k} \neg(x_c \sim k)) \wedge$$

$$(\bigwedge_{\mathfrak{im} \vdash \mathrm{x}_c \sim k} (x_c + \mathfrak{ef}(\pi_0)(c) + \cdots + \mathfrak{ef}(\pi_{\alpha-1})(c) + \mathfrak{ef}(\pi'_\alpha)(c) + \sum_{i,j} z_i^j \mathfrak{ef}(sl_i^j)(c)) \sim k) \wedge$$

$$(\bigwedge_{\text{not } \mathfrak{im} \vdash \mathrm{x}_c \sim k} \neg(x_c + \mathfrak{ef}(\pi_0)(c) + \cdots + \mathfrak{ef}(\pi_{\alpha-1})(c) + \mathfrak{ef}(\pi'_\alpha)(c) + \sum_{i,j} z_i^j \mathfrak{ef}(sl_i^j)(c) \sim k)$$

One more step

- Sequence of small extended paths $\mathbf{P}_1 \cdots \mathbf{P}_{L'}$.
- There is $\varphi(\bar{x}, \bar{y})$ such that

$$[\![\varphi]\!] = \{\langle \mathbf{x}, \mathbf{y} \rangle : \text{ there is a run } \langle q_0, \mathbf{x} \rangle \xrightarrow{*} \langle q_f, \mathbf{y} \rangle \text{ respecting } \mathbf{P}_1 \cdots \mathbf{P}_{L'}\}$$

- $\varphi_i(\bar{x}, \bar{y})$ for each $\mathbf{P}_i$.

$$\exists\, \bar{z_0}, \ldots, \bar{z_{L'}}\ (\bar{x} = \bar{z_0}) \wedge (\bar{y} = \bar{z_{L'}}) \wedge$$

$$\varphi_1(\bar{z_0}, \bar{z_1}) \wedge \varphi_2(\bar{z_1}, \bar{z_2}) \wedge \cdots\ \varphi_{L'-1}(\bar{z_{L'-2}}, \bar{z_{L'-1}}) \wedge \varphi_{L'}(\bar{z_{L'-1}}, \bar{z_{L'}}).$$

- r-reversal-bounded $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$ that is for some $r \geq 0$.
- For each $q' \in Q$, the set

  $$\{\mathbf{y} \in \mathbb{N}^d : \langle q, \mathbf{x} \rangle \xrightarrow{*} \langle q', \mathbf{y} \rangle\}$$

  is a computable Presburger set.
- Formula $\varphi(\bar{\mathrm{y}})$:

  $$\exists\, \bar{\mathrm{x}} \; (\bigwedge_{i \in [1,d]} \mathbf{x}(i) = \mathrm{x}_i) \wedge \bigvee_{\text{small seq. } \sigma = \mathbf{P}_1 \cdots \mathbf{P}_{L'} \text{ ending by } q'} \varphi_{\sigma}(\bar{\mathrm{x}}, \bar{\mathrm{y}})$$

- Assuming that $\mathcal{M}$ is uniformly r-reversal-bounded for some $r \geq 0$. For all q, q', one can compute $\varphi(\bar{\mathrm{x}}, \bar{\mathrm{y}})$ such that

  $$[\![\varphi]\!] = \{\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{N}^{2d} : \langle q, \mathbf{x} \rangle \xrightarrow{*} \langle q', \mathbf{y} \rangle\}$$

Time to reap the rewards!

- Reachability problem with bounded number of reversals.

  Input: a CM $\mathcal{M}$, $r \in \mathbb{N}$, $\langle q_0, \mathbf{x}_0 \rangle$ and $\langle q_f, \mathbf{x}_f \rangle$.

  Question: Is there a run from $\langle q_0, \mathbf{x}_0 \rangle$ to $\langle q_f, \mathbf{x}_f \rangle$ such that each counter has at most r reversals?

- When $\langle \mathcal{M}, \langle q_0, \mathbf{x}_0 \rangle \rangle$ is r'-reversal-bounded for some $r' \leq r$, we get an instance of the reachability problem with initial configuration $\langle q_0, \mathbf{x_0} \rangle$.

- The reachability problem with bounded number of reversals is decidable.

Complexity

- The reachability problem with bounded number of reversals is NP-complete, assuming that all the natural numbers are encoded in binary except the number of reversals.
- The problem is NEXPTIME-complete assuming that all the natural numbers are encoded in binary.

  [Gurari & Ibarra, ICALP'81; Howell & Rosier, JCSS 87]
- NEXPTIME-hardness as a consequence of the standard simulation of Turing machines. [Minsky, 67]

Two or Three Extensions

Adding equality constraints

- Guards so far:

$$g ::= \top \mid \bot \mid \mathrm{x} \sim k \mid g \wedge g \mid g \vee g \mid \neg g$$

  where $\sim \in \{\leq, \geq, =\}$ and $k \in \mathbb{N}$.

- Adding equalities $\mathrm{x} = \mathrm{x}'$ and inequalities $\mathrm{x} \neq \mathrm{x}'$.
- Updates are still equal to $\mathbf{a} \in \mathbb{Z}^d$.

Deterministic Minsky machines

- A counter stores a single natural number.
- A Minsky machine can be viewed as a finite-state machine with two counters.
- Operations on counters:
  - Check whether the counter is zero.
  - Increment the counter by one.
  - Decrement the counter by one if nonzero.

2-counter Minsky machines

- Set of n instructions.
- The lth instruction has one of the forms below ($i \in \{1, 2\}$, $l' \in \{1, \ldots, n\}$):

$$\begin{aligned} \text{l: } & \mathrm{x}_i := \mathrm{x}_i + 1;\text{ goto } l' \\ \text{l: } & \text{if } \mathrm{x}_i = 0 \text{ then goto } l' \text{ else } \mathrm{x}_i := \mathrm{x}_i - 1;\text{ goto } l'' \\ \text{n: } & \text{halt} \end{aligned}$$

- Configurations are elements of $[1, n] \times \mathbb{N} \times \mathbb{N}$.
- Initial configuration: $\langle 1, 0, 0 \rangle$.

Computations

- A computation is a sequence of configurations starting from the initial configuration and such that two successive configurations respect the instructions.

- The Minsky machine

  1: $\mathtt{x}_1 := \mathtt{x}_1 + 1$; goto 2
  2: $\mathtt{x}_2 := \mathtt{x}_2 + 1$; goto 1
  3: halt

  has unique computation

  $\langle 1, 0, 0\rangle \to \langle 2, 1, 0\rangle \to \langle 1, 1, 1\rangle \to \langle 2, 2, 1\rangle \to \langle 1, 2, 2\rangle \to \langle 2, 3, 2\rangle \ldots$

Halting problem

- Halting problem:

  input: a 2-counter Minsky machine $\mathcal{M}$;
  question: is there a finite computation that ends with location equal to n?

  (n is understood as a special instruction that halts the machine)

- Theorem: The halting problem is undecidable. [Minsky,67]

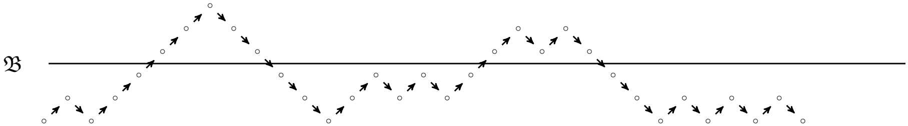
- Minsky machines are Turing-complete.

Undecidability

- Minsky machine $\mathcal{M}$ with *n* instructions and 2 counters.
- Each counter x in $\mathcal{M}$ is given two counters x^{inc} and x^{dec}.
- Zero-test on x is simulated by the guard $\mathrm{x}^{inc} = \mathrm{x}^{dec}$.
- A decrement on x first check that $\mathrm{x}^{inc} \neq \mathrm{x}^{dec}$ and then increment x^{dec}.
- $\mathcal{M}$ can be simulated by a 0-reversal-bounded counter machine with four counters.
- $\mathcal{M}$ halts iff the set of counter values for reaching the state *n* in the 0-reversal-bounded counter machine is not empty.

Weak reversal-boundedness

- Reversals are recorded only above a bound $\mathfrak{B}$:



- Effective semilinearity of the reachability sets.

[Finkel & Sangnier, MFCS'08]

Formal definition

- Counter machine $\mathcal{M} = \langle Q, T, C \rangle$ and bound $\mathfrak{B} \in \mathbb{N}$.
- From $\rho = \langle q_0, \mathbf{x}_0 \rangle \xrightarrow{t_1} \langle q_1, \mathbf{x}_1 \rangle, \ldots$, we defined a sequence of mode vectors $\mathfrak{md}_0, \mathfrak{md}_1, \ldots$ with each $\mathfrak{md}_i \in \{\mathrm{INC}, \mathrm{DEC}\}^d$.
- Set of positions $Rev_i^{\mathfrak{B}}$:

  $$\{j \in [0, |\rho| - 1] : \mathfrak{md}_j(i) \neq \mathfrak{md}_{j+1}(i), \{\mathbf{x}_j(i), \mathbf{x}_{j+1}(i)\} \not\subseteq [0, \mathfrak{B}]\}$$

- $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$ is r-reversal-$\mathfrak{B}$-bounded $\overset{\text{def}}{\Leftrightarrow}$ for every finite run ρ starting at $\langle q, \mathbf{x} \rangle$, $\mathrm{card}(Rev_i^{\mathfrak{B}}) \leq r$ for every $i \in [1, d]$.
- $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$ is weakly reversal-bounded $\overset{\text{def}}{\Leftrightarrow}$ there are $r, \mathfrak{B} \geq 0$ such that $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$ is r-reversal-$\mathfrak{B}$-bounded.
- r-reversal-boundedness = r-reversal-0-boundedness.

Reachability sets are Presburger sets too!

- r-reversal-$\mathfrak{B}$-bounded counter machine $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$.
- For each $q' \in Q$,

  $$\{\mathbf{y} \in \mathbb{N}^d : \langle q, \mathbf{x} \rangle \xrightarrow{*} \langle q', \mathbf{y} \rangle\}$$

  is a computable Presburger set.
- This extends the results for r-reversal-boundedness.
- …but the proof uses simply those results.

$$\boxed{\textbf{The Reversal-Boundedness Detection Problem}}$$

The reversal-boundedness detection problem

- The reversal-boundedness detection problem:

  Input: Counter machine $\mathcal{M}$ of dimension d, configuration $\langle \mathcal{M}, \langle q_0, \mathbf{x}_0 \rangle \rangle$ and $i \in [1, d]$.

  Question: Is $\langle \mathcal{M}, \langle q_0, \mathbf{x}_0 \rangle \rangle$ reversal-bounded with respect to the counter $\mathtt{x}_i$?

- Undecidability due to [Ibarra, JACM 78].
- Restriction to VASS is decidable [Finkel & Sangnier, MFCS'08].

Undecidability proof

- Minsky machine $\mathcal{M}$ with halting state q_H (2 counters).
- Either $\mathcal{M}$ has a unique infinite run (and never visits q_H) or $\mathcal{M}$ has a finite run (and halts at q_H).
- Counter machine $\mathcal{M}'$: replace $t = q_i \xrightarrow{\varphi} q_j$ by

$$q_i \xrightarrow{\mathtt{++x_1}} q_{1,t}^{new} \xrightarrow{\mathtt{--x_1}} q_{2,t}^{new} \xrightarrow{\varphi} q_j$$

- We have the following equivalences:
  - $\mathcal{M}$ halts.
  - For $\mathcal{M}'$, q_H is reached from $\langle q_0, \mathbf{0} \rangle$.
  - Unique run of $\mathcal{M}'$ starting by $\langle q_0, \mathbf{0} \rangle$ is finite.
  - $\mathcal{M}'$ is reversal-bounded from $\langle q_0, \mathbf{0} \rangle$.

Decidable Repeated Reachability Problems

The problems

- Control state repeated reachability problem with bounded number of reversals:

  Input: CM $\mathcal{M}$, $\langle q_0, \mathbf{x}_0 \rangle$, $r \geq 0$, state q_f.

  Question: is there an infinite r-reversal-bounded run starting from $\langle q_0, \mathbf{x}_0 \rangle$ such that q_f is repeated infinitely often?

- Control state reachability problem with bounded number of reversals:

  Input: CM $\mathcal{M}$, $\langle q_0, \mathbf{x}_0 \rangle$, $r \geq 0$, state q_f.

  Question: is there a finite r-reversal-bounded run starting from $\langle q_0, \mathbf{x}_0 \rangle$ such that q_f is reached?

- Control state reachability problem with bounded number of reversals is decidable.

- Control state repeated reachability problem with bounded number of reversals is decidable.

[Dang & Ibarra & San Pietro, FSTTCS'01]

Next lecture on October 14th

- Lecturer: Alain Finkel (finkel@lsv.fr).

Exercises

- Show that the class of ultimately periodic sets is closed under union and intersection.
- Show that for every linear set there is an initialized 0-reversal-bounded counter machine whose reachability set is equal to it.

Exercise (1/5)

- Goal: Show decidability of the problem:

  Input: $\langle \mathcal{M}, \langle q, \mathbf{x} \rangle \rangle$ and semilinear set $X \subseteq \mathbb{N}^d$ defined by $\langle \mathbf{b}_1, \mathfrak{P}_1 \rangle, \ldots, \langle \mathbf{b}_\alpha, \mathfrak{P}_\alpha \rangle$.

  Question: Is there an infinite r-reversal-bounded run from $\langle q, \mathbf{x} \rangle$ such that infinitely often the counter values are in X?

A) Show that we can restrict ourselves to $\alpha = 1$ and infinitely often the counter values belong to the linear set $\langle \mathbf{b}_1, \mathfrak{P}_1 \rangle$ and simulaneously the location is some fixed q'.

Exercise (2/5)

B) Linear set X characterised by $\mathbf{b}$ and $\mathbf{p}_1, \ldots, \mathbf{p}_N$.
Let $\mathbf{x}_1, \mathbf{x}_2, \ldots$ be an infinite sequence of elements in X.
Show that there are $\ell' < \ell$ and $\mathbf{a}, \mathbf{c} \in \mathbb{N}^N$ such that

(I) $\mathbf{x}_{\ell'} \preceq \mathbf{x}_{\ell}$,

(II) $\mathbf{x}_{\ell'} = \mathbf{b} + \sum_{k \in [1,N]} \mathbf{a}(k)\mathbf{p}_k$,

(III) $\mathbf{x}_{\ell} = \mathbf{b} + \sum_{k \in [1,N]} \mathbf{c}(k)\mathbf{p}_k$,

(IV) $\mathbf{a} \preceq \mathbf{c}$.

C) Design a 0-reversal-bounded counter machine with d counters such that for some state $q_0, q_f \in Q$, for all $\mathbf{x} \in \mathbb{N}^d$, $\mathbf{x} \in X$ iff there is a run from $\langle q_0, \mathbf{x} \rangle$ to $\langle q_f, \mathbf{0} \rangle$.

Exercise (3/5)

D) Design a 1-reversal-bounded CM with $2d$ counters such that for some state $q_0, q_f \in Q$, for all $\mathbf{x} \in \mathbb{N}^{2d}$ such that the restriction to $\mathbf{x}$ to the d last counters equal to $\mathbf{0}$,

the restriction of $\mathbf{x}$ to the d first counters belongs to X
iff
there is a run from $\langle q_0, \mathbf{x} \rangle$ to $\langle q_f, \mathbf{x} \rangle$.

E) Design a 1-reversal-bounded CM with $4d$ counters such that for some state $q_0, q_f \in Q$, for all $\mathbf{x} \in \mathbb{N}^{4d}$ such that the restriction to $\mathbf{x}$ to the $2d$ last counters equal to $\mathbf{0}$,

there are $\lambda_1, \ldots, \lambda_N \in \mathbb{N}$ such that for all $i \in [1, d]$,
$$\mathbf{x}(d+i) - \mathbf{x}(i) = \lambda_1 \mathbf{p}_1(i) + \cdots \lambda_N \mathbf{p}_N(i)$$
iff
there is a run from $\langle q_0, \mathbf{x} \rangle$ to $\langle q_f, \mathbf{x} \rangle$.

Exercise (4/5)

Show that the conditions below are equivalent:

($\star$) There is an infinite r-reversal-bounded run from $\langle q_0, \mathbf{x}_0 \rangle$ such that counter values belong to X and the state is q' infinitely often.

($\star\star$) There exist a finite r-reversal-bounded run $\rho = \langle q_0, \mathbf{x}_0 \rangle \xrightarrow{t_1} \langle q_1, \mathbf{x}_1 \rangle \cdots \xrightarrow{t_l} \langle q_\ell, \mathbf{x}_\ell \rangle$, $\ell' \in [0, \ell - 1]$ and $C_= \subseteq C$ such that

(a) $q_\ell = q_{\ell'} = q'$,
(b) $\mathbf{x}_{\ell'}, \mathbf{x}_\ell \in X$,
(c) (I)–(IV) above,
(d) for $\mathrm{x}_i \in C_=$ and $j \in [\ell' + 1, \ell]$, $\mathbf{x}_j(i) - \mathbf{x}_{j-1}(i) = 0$,
(e) for $\mathrm{x}_i \in (C \setminus C_=)$ and $j \in [\ell' + 1, \ell]$, $\mathbf{x}_{j-1}(i) \leq \mathbf{x}_j(i)$,
(f) for $\mathrm{x}_i \in (C \setminus C_=)$, we have $k_{max} < \mathbf{x}_{\ell'}(i)$.
(g) for all $\mathrm{x}_i \in C_=$, have $\mathbf{x}_{\ell'}(i) \leq k_{max}$.

k_{max}: maximal constant k occurring in guards

Exercise (5/5)

- Design a reduction from $(\star\star)$ to an instance of the reachability problem with bounded number of reversals.
- Conclude that checking whether an initialized counter machine has an infinite r-reversal-bounded run visiting infinitely often a semilinear set can be decided in NEXPTIME.