Presburger Arithmetic

- Which arithmetical problems can be solved using automata?

- Presburger arithmetic (PA): a logical language to define arithmetical properties of (tuples of) natural numbers

Is there an integer solution?

$$3x - 4y = 5$$
$$-x + y = 3$$

Is there an integer solution?

$$2x + 3y \geq 5$$
$$-x + 4y \leq 3$$

Are there integers x, y such that

$$3x - 4y = 5$$
$$-x + y = 3$$

but not

$$2x + 3y \geq 2$$
$$-x + 4y \leq 4$$

?

For every integer solution x, y of

$$2x + 3y \geq 5$$

$$-x + 4y \leq 3$$

is there is an integer solution z, u of

$$3z - 2u \geq 3$$

$$-z + 4u \leq -2$$

such that $x + z = y + u$?

Syntax of PA

- Symbols:

Variables	x, y, z ...
Constants	0, 1
Arithmetical symbols	$+$, $\leq$
Logical symbols	$\vee$, $\neg$, $\exists$ ($\wedge$, $\forall$, $\rightarrow$,...)
Parenthesis	(,)

- Terms:

  Variables, 0 and 1 are terms.

  If t and u are terms, then $t + u$ is a term.

Syntax of PA

- Atomic formulas:

  $t \leq u$, where t and u are terms

- Formulas:

  Atomic formulas are formulas.

  If φ_1, φ_2 are formulas, then so are $\varphi_1 \vee \varphi_2, \neg \varphi_1, \exists x \varphi_1$

- Free and bound variables:

  A variable is bound if it is in the scope of an existential quantifier, otherwise it is free.

- Sentences: formulas without free variables.

Abbreviations

- Logical abbrevations:

$$\varphi_1 \wedge \varphi_2 \equiv \neg(\neg\varphi_1 \vee \neg\varphi_2)$$

$$\varphi_1 \rightarrow \varphi_2 \equiv \neg\varphi_1 \vee \varphi_2$$

$$\varphi_1 \leftrightarrow \varphi_2 \equiv \neg(\varphi_1 \vee \varphi_2) \vee \neg(\neg\varphi_1 \vee \neg \varphi_2)$$

$$\forall x\, \varphi \equiv \neg \exists x \neg \varphi$$

- Arithmetic abbreviations:

$$n := \underbrace{1 + 1 + \ldots + 1}_{n \text{ times}} \qquad nx := \underbrace{x + x + \ldots + x}_{n \text{ times}}$$

$$\begin{aligned} t \geq t' &:= t' \leq t \\ t = t' &:= t \leq t' \wedge t \geq t' \\ t < t' &:= t \leq t' \wedge \neg(t = t') \\ t > t' &:= t' < t \end{aligned}$$

Semantics (intuition)

- The semantics of a sentence is true or false.
- The semantics of a formula with free variables $(x_1, \ldots, x_k)$ is the set containing all tuples $(n_1, \ldots, n_k)$ of natural numbers that ‘’satisfy the formula’’

Semantics (more formally)

- An interpretation of a formula φ is a function $\mathcal{I}$ that assigns a natural number to every free variable appearing in φ (and perhaps also to others).
- Given an interpretation $\mathcal{I}$, a variable x, and a number n, we denote by $\mathcal{I}[n/x]$ the interpretation that assigns to x the number n, and to all other variables the same value as $\mathcal{I}$.

Semantics (more formally)

- We inductively define when an interpretation $\mathfrak{I}$ satisfies a formula φ, denoted by $\mathfrak{I} \models \varphi$:

$$
\begin{array}{lll}
\mathfrak{I} \models t \leq u & \text{iff} & \mathfrak{I}(t) \leq \mathfrak{I}(u) \\
\mathfrak{I} \models \neg\varphi_1 & \text{iff} & \mathfrak{I} \not\models \varphi_1 \\
\mathfrak{I} \models \varphi_1 \vee \varphi_2 & \text{iff} & \mathfrak{I} \models \varphi_1 \text{ or } \mathfrak{I} \models \varphi_2 \\
\mathfrak{I} \models \exists x\, \varphi & \text{iff} & \text{there exists } n \geq 0 \text{ such that } \mathfrak{I}[n/x] \models \varphi
\end{array}
$$

Semantics (more formally)

- Lemma: If two interpretations of a formula φ assign the same values to all free variables of φ, then either both satisfy φ or none satisfy φ.
- Corollary: if φ is a sentence, either all interpretations satisfy φ, or none satisfy φ.
- A model or solution of φ is the projection of an interpretation that satisfies φ onto the free variables of φ. The set of solutions or solution space is denoted by $Sol(\varphi)$.

Formulating questions

Are there integers x, y such that

$$2x + 3y \geq 5$$

$$-x + 4y \leq 3 \quad ?$$

$$\exists x \exists y \quad (2x + 3y \geq 5 \wedge -x + 4y \leq 3)$$

Formulating questions

For every solution x, y of

$$2x + 3y \geq 5$$

$$-x + 4y \leq 3$$

is there is a solution z, u of

$$3z - 2u \geq 3$$

$$-z + 4u \leq -2$$

such that $x + z = y + u$?

$$\forall x \forall y$$

$$(2x + 3y \geq 5 \wedge -x + 4y \leq 3)$$

$$\rightarrow$$

$$(\exists z \exists u$$

$$(3z - 2u \geq 3 \wedge$$

$$-z + 4u \leq -2 \wedge$$

$$x + z = y + u))$$

Language of a formula

- We encode natural numbers with the $lsbf$ encoding.
- If φ has free variables $x_1, \ldots, x_k$, we encode a solution of φ as a word over $\{0,1\}^k$ in the usual way. E.g, the minimal encoding of $(x_1, x_2, x_3) = (5,10,0)$ is

$$\begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

- The language of φ, denoted by $L(\varphi)$, is the set of encodings of the solutions of φ.

An NFA for the solution space

- Given φ, we construct an NFA A_φ such that $L(A_\varphi) = L(\varphi)$
- We can take:

$$A_{\neg\varphi} \quad := \quad CompNFA(A_\varphi)$$

$$A_{(\varphi_1 \vee \varphi_2)} \quad := \quad UnionNFA(A_{\varphi_1}, A_{\varphi_2})$$

$$A_{\exists x \varphi} \quad := \quad Projection\_x(A_\varphi)$$

where $Projection\_x$ projects onto all variables but x

- It remains to construct A_φ for an atomic formula φ.

DFA for atomic formulas

- Every atomic formula has the same solutions as an equation of the form

$$a_1 x_1 + \ \dots + a_n x_n \ \leq b \ \coloneqq a \cdot x \leq b$$

where the a_i and b are arbitrary integers (possibly negative).

- Given $a \cdot x \leq b$ we construct a DFA with integers as states and b as initial state satisfying:

Each state $q \in \mathbb{Z}$ recognizes the tuples $c \in \mathbb{N}^n$ such that $a \cdot c \leq q$

Transitions

- Given $q \in \mathbb{Z}$ and a letter $\zeta \in \{0,1\}^n$ we compute the target state $q' \in \mathbb{Z}$ of the transition (q, ζ, q').
- For every word $w \in (\{0,1\}^n)^*$ we have:

$$w \text{ is accepted from } q' \text{ iff } \zeta w \text{ is accepted from } q$$

and so for every tuple $c \in \mathbb{N}^n$:

$$c \text{ is accepted from } q' \text{ iff } 2c + \zeta \text{ is accepted from } q$$

- Hence we choose q' so that

$$a \cdot c \leq q' \text{ iff } a \cdot (2c + \zeta) \leq q$$

- Since $a \cdot (2c + \zeta) \leq q$ iff $2(a \cdot c) + a \cdot \zeta \leq q$ iff $a \cdot c \leq \left\lfloor \frac{q - a \cdot \zeta}{2} \right\rfloor$ we take

$$q' = \left\lfloor \frac{q - a \cdot \zeta}{2} \right\rfloor$$

Final states

- A state is final iff it accepts the empty word
- So $q \in \mathbb{Z}$ is final iff it accepts $(0, \ldots, 0) \in \mathbb{N}^n$
- So we take $q \in \mathbb{Z}$ final iff $a \cdot (0, \ldots, 0) \leq q$ iff $q \geq 0$

AFtoDFA(φ)

Input: Atomic formula $\varphi = a \cdot x \leq b$

Output: DFA $A_\varphi = (Q, \Sigma, \delta, q_0, F)$ such that $L\left(A_\varphi\right) = L(\varphi)$

$Q, \delta, F \leftarrow \emptyset$; $q_0 \leftarrow s_b$
$W \leftarrow \{s_b\}$
while $W \neq \emptyset$ **do**
    pick s_k **from** W
    add s_k **to** Q
    if $k \geq 0$ **then add** s_k **to** F
    for all $\zeta \in \{0, 1\}^n$ **do**
        $j \leftarrow \left\lfloor \frac{k - a \cdot \zeta}{2} \right\rfloor$
        if $s_j \notin Q$ **then add** s_j **to** W
        add (s_k, ζ, s_j) **to** δ

Example: $3x - 2y \geq 6$

Conversion: $-3x + 2y \leq -6$

$$a = \binom{-3}{2}, \quad b = -6$$

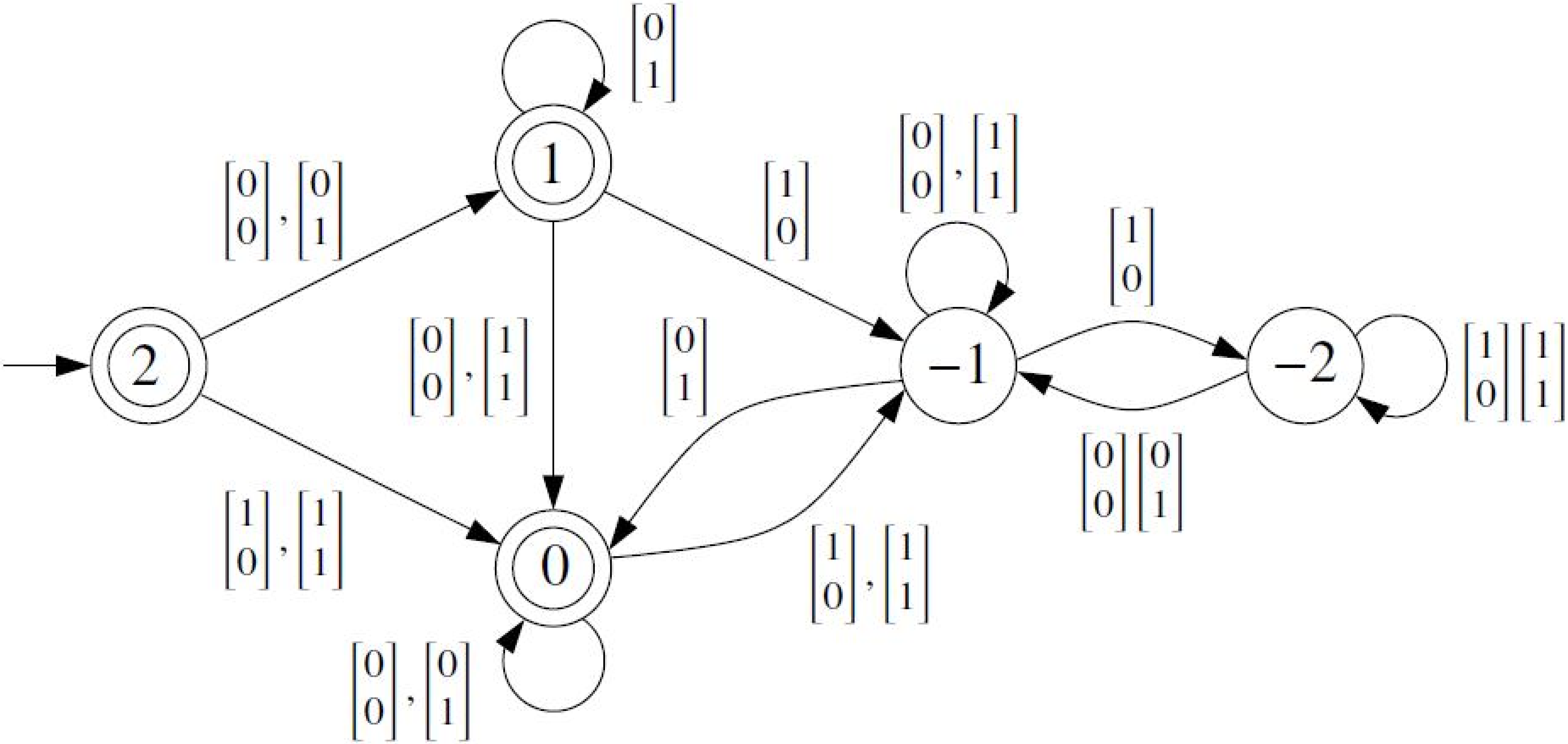
Initial state: -6

Transition from state -6 with letter $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$:

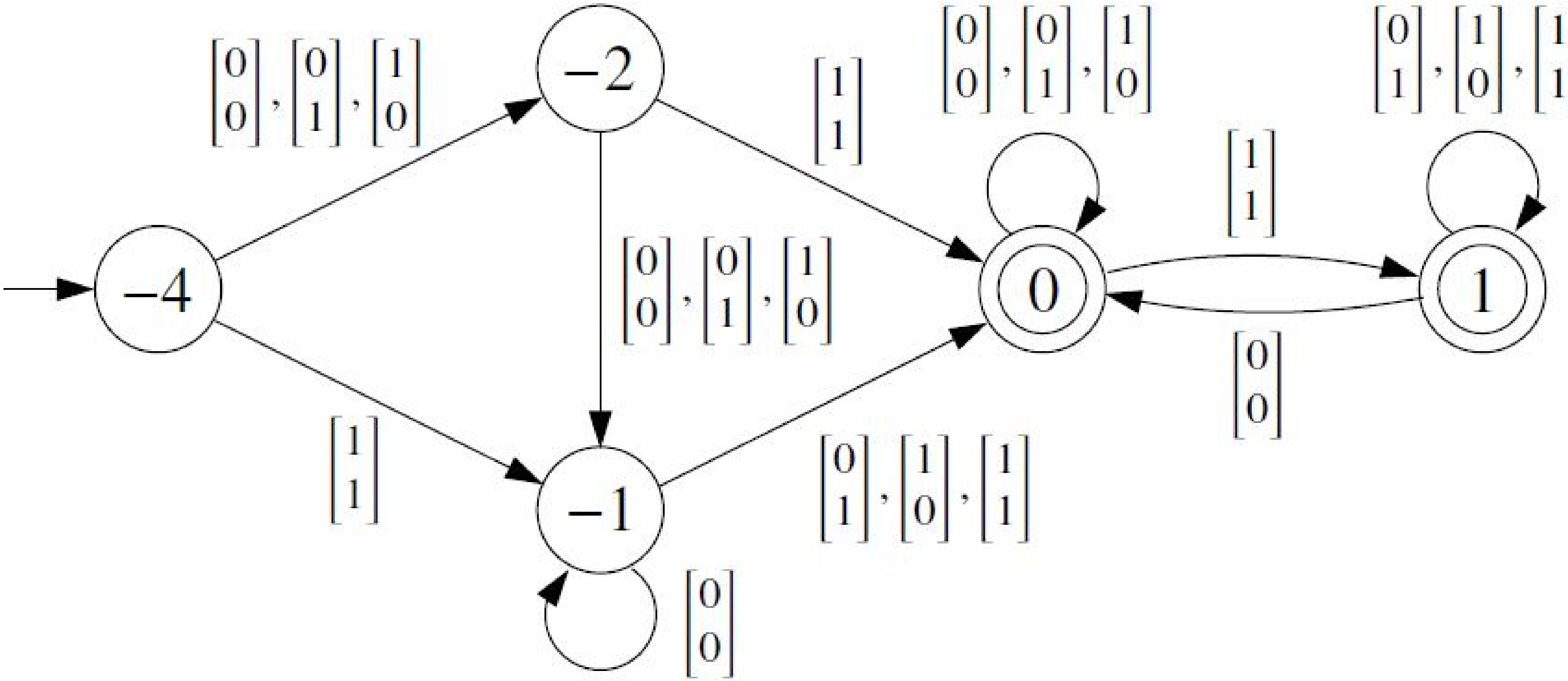
$$q' = \left\lfloor \frac{q - a \cdot \zeta}{2} \right\rfloor$$

$$q' = \left\lfloor \frac{-6 - (-3,2) \cdot \binom{1}{1}}{2} \right\rfloor = \left\lfloor \frac{-6 + 1}{2} \right\rfloor = -3$$

Example: $2x - y \leq 2$



Example: $x + y \geq 4$



Termination of *AFtoDFA*

- Lemma: Let $\varphi = a \cdot c \leq b$ and $s = \sum_{i=1}^{n} |a_i|$.
  All states s_j added by *AFtoDFA*(φ) satisfy
  $$-|b| - s \leq j \leq |b| + s$$
  Proof: Holds for the first state added: s_b

  Assume s_j is added to the workset when processing s_k.
  By ind. hyp.: $-|b| - s \leq k \leq |b| + s$.

  Together with $j = \left\lfloor \frac{k - a \cdot \zeta}{2} \right\rfloor$ we get
  $$\left\lfloor \frac{-|b| - s - a \cdot \zeta}{2} \right\rfloor \leq j \leq \left\lfloor \frac{|b| + s - a \cdot \zeta}{2} \right\rfloor$$

$$\left\lfloor \frac{-|b| - s - a \cdot \zeta}{2} \right\rfloor \leq j \leq \left\lfloor \frac{|b| + s - a \cdot \zeta}{2} \right\rfloor$$

Some arithmetic yields

$$-|b| - s \quad \leq \quad \frac{-|b| - 2s}{2} \quad \leq \quad \left\lfloor \frac{-|b| - s - a \cdot \zeta}{2} \right\rfloor$$

$$\left\lfloor \frac{|b| + s - a \cdot \zeta}{2} \right\rfloor \quad \leq \quad \frac{|b| + 2s}{2} \quad \leq \quad |b| + s$$

and together we get

$$-|b| - s \leq j \leq |b| + s$$

Solving a system of inequations

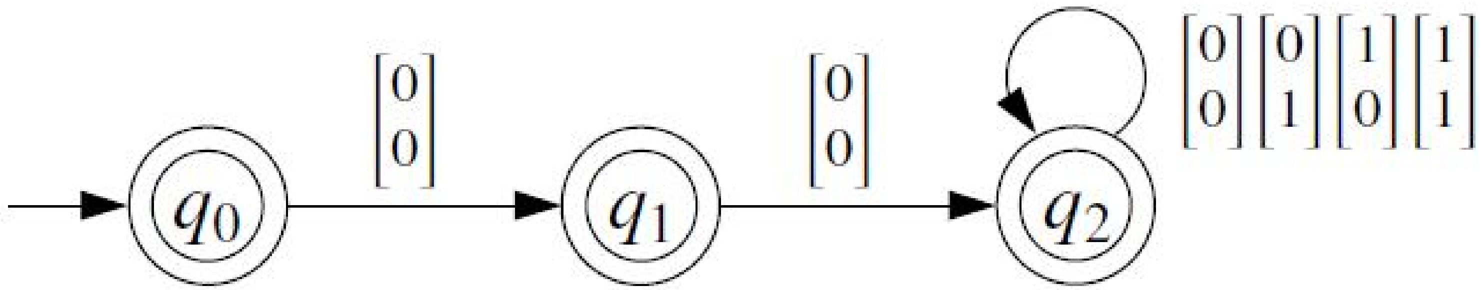
- We compute all solutions of

$$2x - y \leq 2$$
$$x + y \geq 2$$

s.t. x, y are multiples of 4. They are the solutions of

$$(\exists z\ x = 4z) \wedge (\exists w\ y = 4w) \wedge (2x - y \leq 2) \wedge (x + y \geq 4)$$

- DFA for $(\exists z\ x = 4z) \wedge (\exists w\ y = 4w)$



- Final result