

Weighted first-order logic, weighted aperiodic automata, and more

Paul Gastin

LMF, ENS Paris-Saclay, CNRS, Univ. Paris-Saclay, IRL ReLaX

Manfred Droste & P.G., Aperiodic Weighted Automata and Weighted First-Order Logic, MFCS'19
+ ongoing work

Words — ~~Boolean~~

Weighted

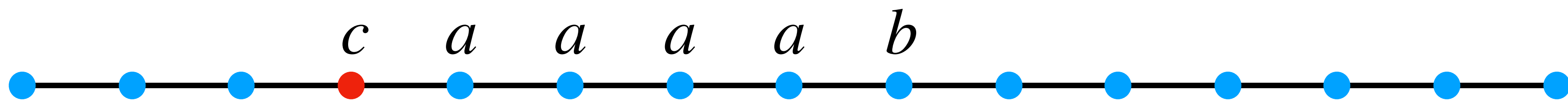
TFAE:

1. First-order logic (FO)
2. Linear Temporal Logic (LTL)
3. Aperiodic Languages (or Automata) (AP)
4. Star-Free Languages (SF)
5. Star-Free Propositional Dynamic Logic (SFPDL)

FO=LTL Kamp PhD 1968

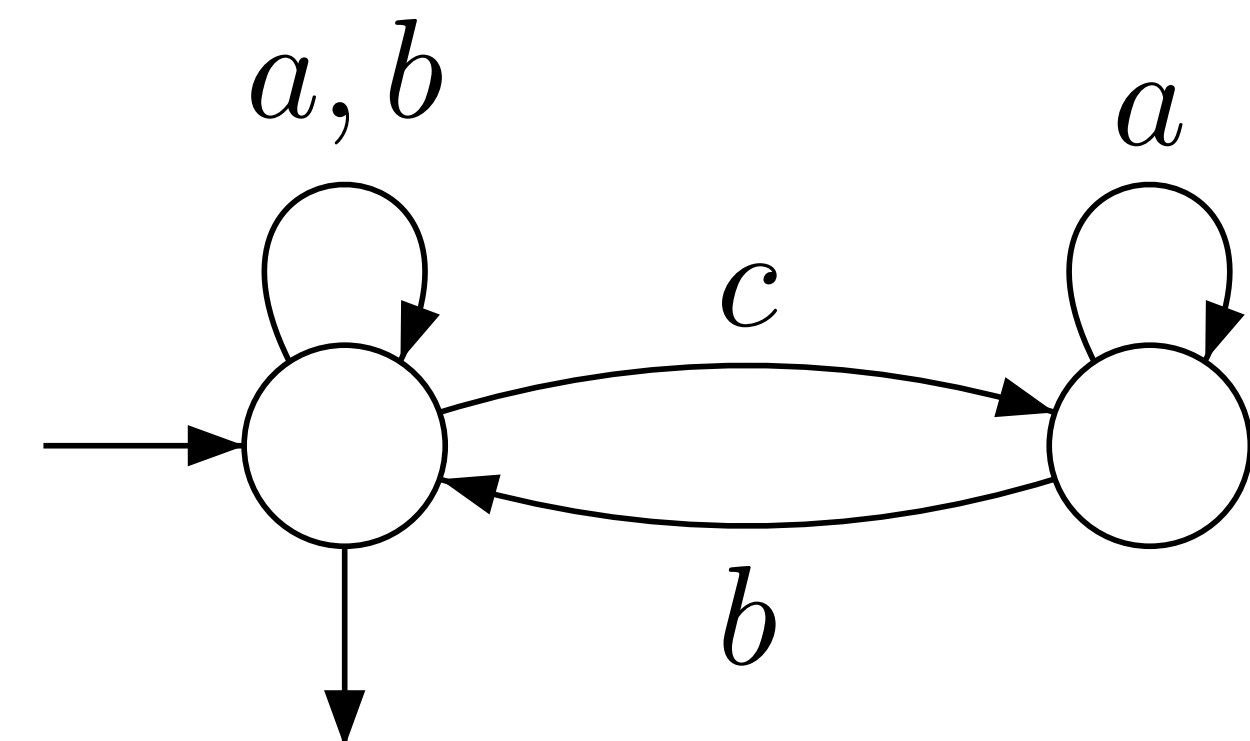
AP=SF Schützenberger I&C 1965

SF=FO McNaughton & Papert 1971

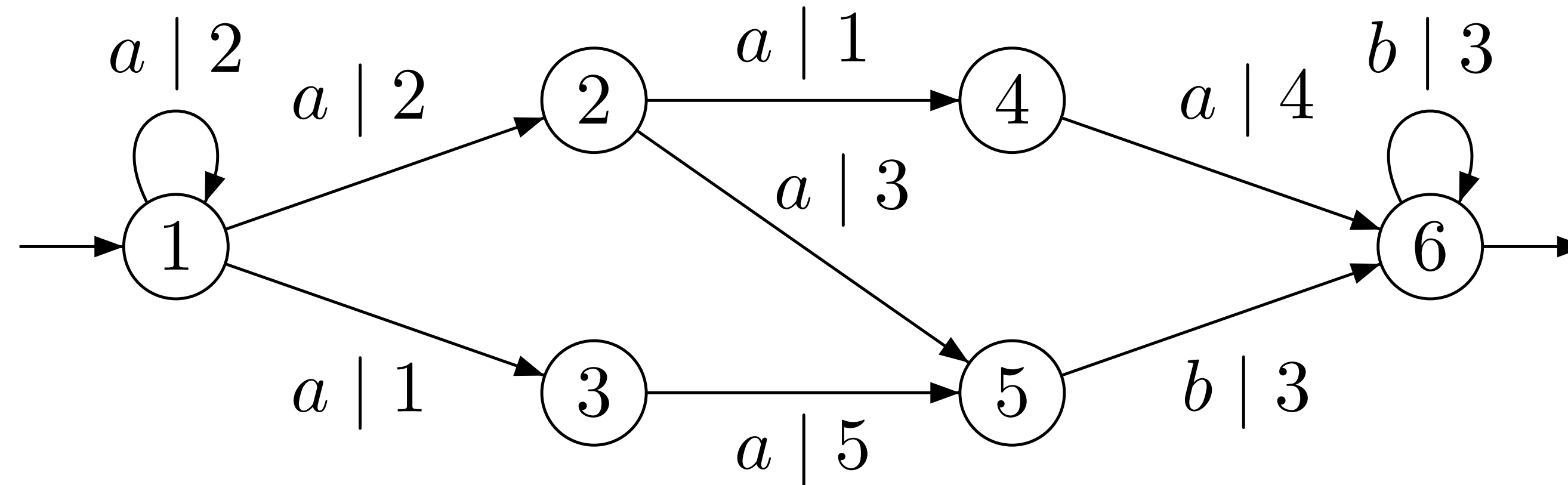


$$\forall x \left(c(x) \rightarrow \exists z \left(z > x \wedge b(z) \wedge \forall y \left(x < y < z \rightarrow a(y) \right) \right) \right)$$

$$G(c \rightarrow X(a \cup b))$$



Weighted Automata (wA)



Input word: $w = a^n a^3 b b^p$

- $1 \xrightarrow{a} 1 \dots 1 \xrightarrow{a} 2 \xrightarrow{a} 4 \xrightarrow{a} 6 \xrightarrow{b} 6 \dots 6 \xrightarrow{b} 6$
- $1 \xrightarrow{a} 1 \dots 1 \xrightarrow{a} 1 \xrightarrow{a} 2 \xrightarrow{a} 5 \xrightarrow{b} 6 \dots 6 \xrightarrow{b} 6$
- $1 \xrightarrow{a} 1 \dots 1 \xrightarrow{a} 1 \xrightarrow{a} 3 \xrightarrow{a} 5 \xrightarrow{b} 6 \dots 6 \xrightarrow{b} 6$

Weights of runs:

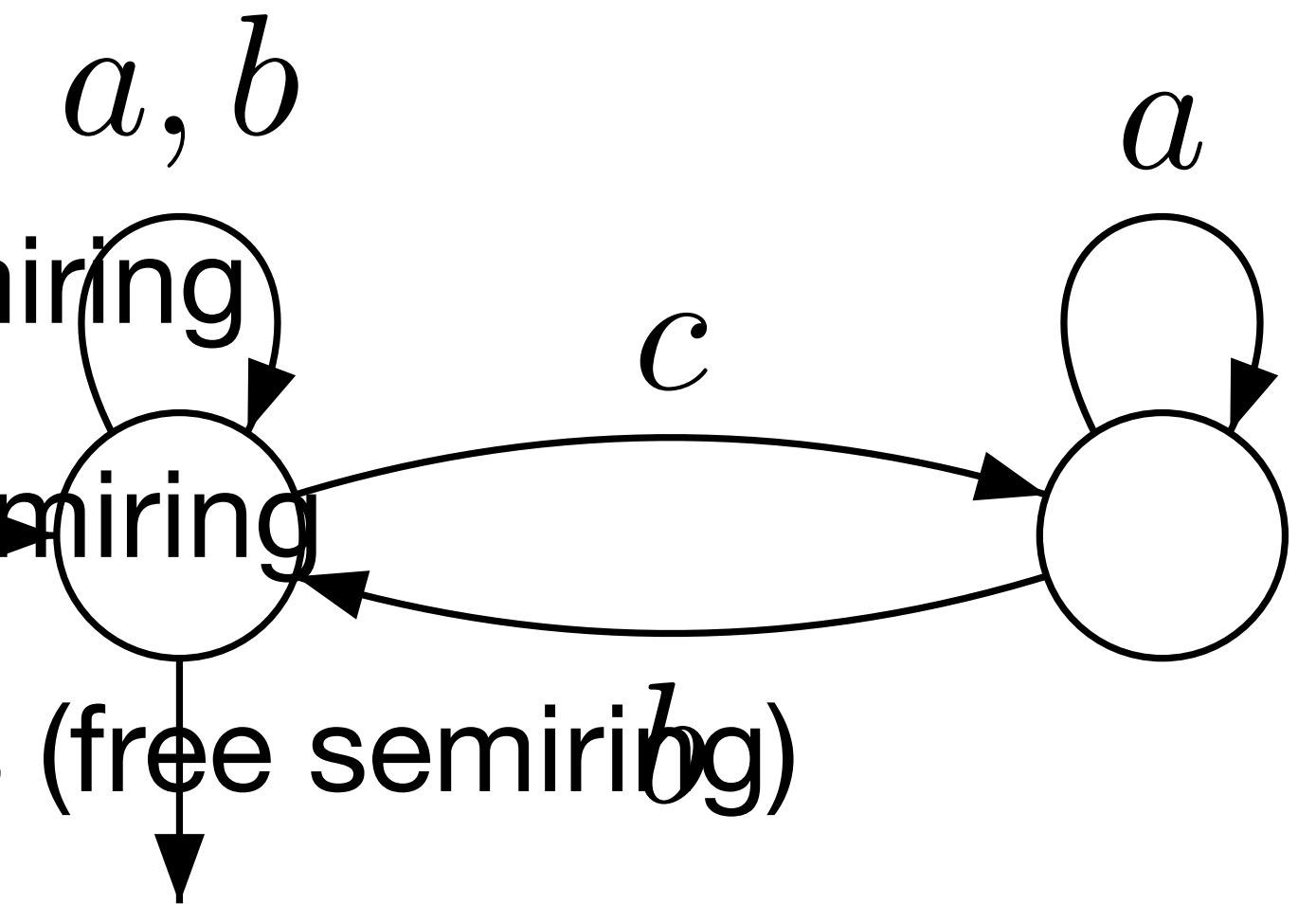
- $2^n \times 2 \times 1 \times 4 \times 3 \times 3^p$
- $2^n \times 2 \times 2 \times 3 \times 3 \times 3^p$
- $2^n \times 2 \times 1 \times 5 \times 3 \times 3^p$

$$[\mathcal{A}](w) = \sum_{\rho} \text{wt}(\rho) = 2^n \times (24 + 36 + 30) \times 3^p$$

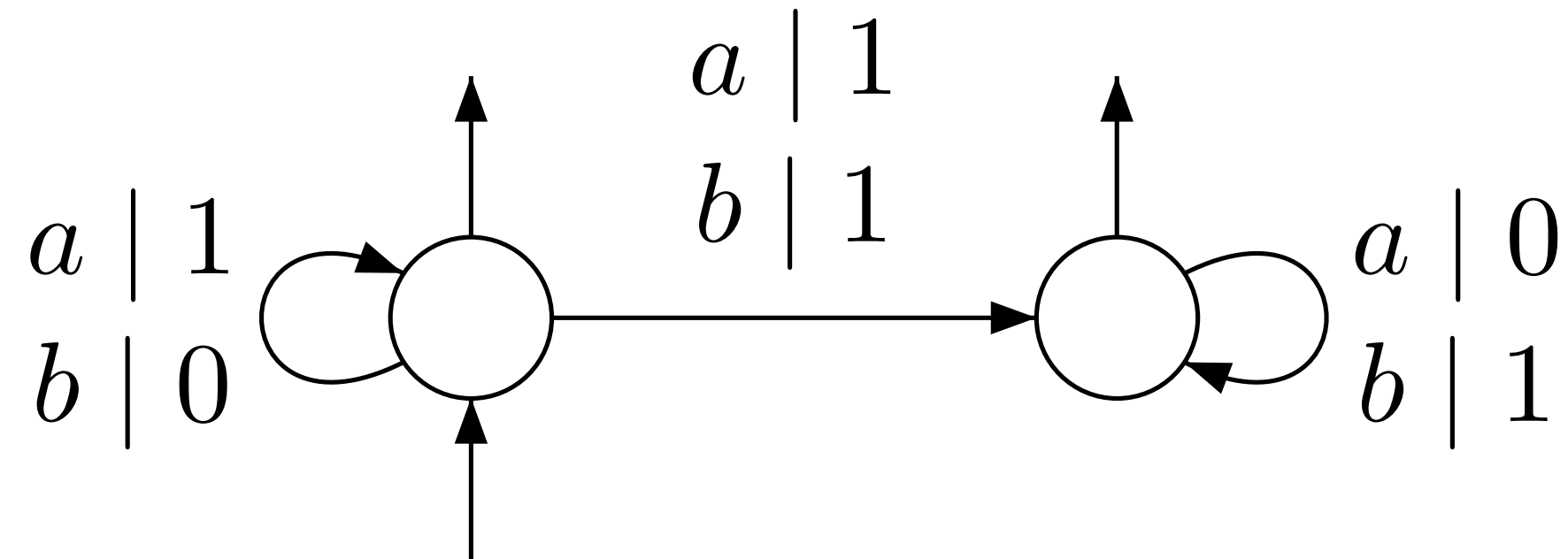
Semirings

Examples:

1. $\mathbb{N}_{+, \times} = (\mathbb{N}, +, \times, 0, 1)$: Natural numbers (usual addition and multiplication)
2. $\mathbb{B} = (\{ \perp, \top \}, \vee, \wedge, \perp, \top)$: Boolean semiring
3. $\mathbb{N}_{\min, +} = (\mathbb{N} \cup \{+\infty\}, \min, +, +\infty, 0)$: min-plus semiring
4. $\mathbb{N}_{\max, +} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$: max-plus semiring
5. $(\mathbb{N}\langle \mathbb{R}^* \rangle, \uplus, \cdot, \emptyset, \{\{\varepsilon\}\})$: multisets of weight sequences (free semiring)



Example with max-plus



$\mathbb{N}_{\max,+} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0) : \text{max-plus semiring}$

$$[[\mathcal{A}]](w) = \max \{ |u|_a + |v|_b \mid w = uv \}$$

Weighted First-Order Logic (wFO)

$$\text{FO} \quad \varphi ::= \top \mid a(x) \mid x \leq y \mid \neg \varphi \mid \varphi \vee \varphi \mid \exists x \varphi$$

$$\text{step-wFO} \quad \alpha ::= r \mid \varphi ? \alpha : \alpha$$

$$\text{wFO} \quad \Phi ::= \mathbf{0} \mid \prod_z \alpha \mid \sum_x \Phi \mid \Phi + \Phi \mid \varphi ? \Phi : \Phi$$

Example: $L = (ab + aab + c)^+$ is FO-definable

- $\varphi_L = \forall x (b(x) \rightarrow a(x-1)) \wedge (a(x) \rightarrow b(x+1) \vee (a(x+1) \wedge b(x+2)))$
- $\alpha(z) = (a(z) \vee c(z)) ? 1 : (a(z-2) ? 3 : 2)$
- $\Phi = \varphi_L ? \prod_z \alpha(z) : \mathbf{0}$
- For $w \in L$, let $n = |w|_{ab}$ and $m = |w|_{aab}$. Then, $[[\Phi]]^w = 3^m \times 2^{n-m}$

Weighted First-Order Logic (wFO)

FO $\varphi ::= \top \mid a(x) \mid x \leq y \mid \neg\varphi \mid \varphi \vee \varphi \mid \exists x\varphi$

step-wFO $\alpha ::= r \mid \varphi ? \alpha : \alpha$

wFO $\Phi ::= \mathbf{0} \mid \prod_z \alpha \mid \sum_x \Phi \mid \Phi + \Phi \mid \varphi ? \Phi : \Phi$

Semantics: $w \in \Sigma^+$ and $\nu: V \rightarrow \text{pos}(w)$

- $\llbracket r \rrbracket_\nu^w = r$ and $\llbracket \varphi ? \alpha_1 : \alpha_2 \rrbracket_\nu^w = \begin{cases} \llbracket \alpha_1 \rrbracket_\nu^w & \text{if } w, \nu \models \varphi \\ \llbracket \alpha_2 \rrbracket_\nu^w & \text{otherwise.} \end{cases}$
- $\llbracket \mathbf{0} \rrbracket_\nu^w = 0$ and $\llbracket \prod_z \alpha \rrbracket_\nu^w = \prod_{j=1}^{|w|} \llbracket \alpha \rrbracket_{\nu[z \mapsto j]}^w$ and $\llbracket \sum_x \Phi \rrbracket_\nu^w = \sum_{j=1}^{|w|} \llbracket \Phi \rrbracket_{\nu[x \mapsto j]}^w$
- $\llbracket \Phi_1 + \Phi_2 \rrbracket_\nu^w = \llbracket \Phi_1 \rrbracket_\nu^w + \llbracket \Phi_2 \rrbracket_\nu^w$ and $\llbracket \varphi ? \Phi_1 : \Phi_2 \rrbracket_\nu^w = \begin{cases} \llbracket \Phi_1 \rrbracket_\nu^w & \text{if } w, \nu \models \varphi \\ \llbracket \Phi_2 \rrbracket_\nu^w & \text{otherwise.} \end{cases}$

Weighted First-Order Logic (wFO)

$$\text{FO} \quad \varphi ::= \top \mid a(x) \mid x \leq y \mid \neg \varphi \mid \varphi \vee \varphi \mid \exists x \varphi$$

$$\text{step-wFO} \quad \alpha ::= r \mid \varphi ? \alpha : \alpha$$

$$\text{wFO} \quad \Phi ::= \mathbf{0} \mid \prod_z \alpha \mid \sum_x \Phi \mid \Phi + \Phi \mid \varphi ? \Phi : \Phi$$

Example: Let $L \subseteq \Sigma^+$ be a language defined by the FO sentence φ .

$$\Phi_{\text{pref}} = \sum_x \varphi^{\leq x} ? \mathbf{1} : \mathbf{0}$$

where $\mathbf{1} = \prod_z 1$ and $\varphi^{\leq x}$ is φ relativized to the prefix $[1, x]$

Number of
prefixes in L

Weighted First-Order Logic (wFO)

$$\text{FO} \quad \varphi ::= \top \mid a(x) \mid x \leq y \mid \neg \varphi \mid \varphi \vee \varphi \mid \exists x \varphi$$

$$\text{step-wFO} \quad \alpha ::= r \mid \varphi ? \alpha : \alpha$$

$$\text{wFO} \quad \Phi ::= \mathbf{0} \mid \prod_z \alpha \mid \sum_x \Phi \mid \Phi + \Phi \mid \varphi ? \Phi : \Phi$$

Example: Let $L \subseteq \Sigma^+$ be a language defined by the FO sentence φ .

$$\Phi_{\text{inf}} = \sum_{x,y} (x \leq y \wedge \varphi^{[x,y]}) ? \mathbf{1} : \mathbf{0}$$

where $\mathbf{1} = \prod_z 1$ and $\varphi^{[x,y]}$ is φ relativized to the infix $[x, y]$

Number of
factors in L

wFO vs wA

Theorem [Manfred Droste & P.G., 2019]

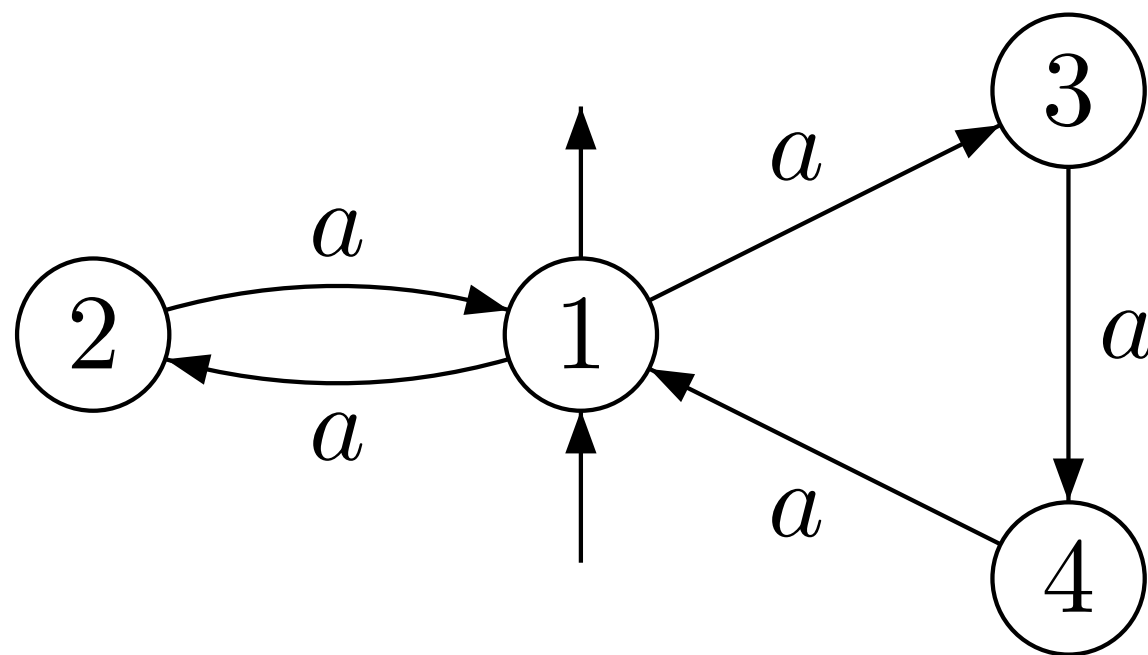
The following classes are expressively equivalent:

- Aperiodic polynomially ambiguous wA and wFO sentences,
- Aperiodic finitely ambiguous wA and wFO sentences without first-order sums ($\sum_x$),
- Aperiodic unambiguous wA and wFO sentences without binary sums (+) or first-order sums ($\sum_x$).

Aperiodic automata

The transition monoid of $\mathcal{A}$ is aperiodic:

$$\exists n \geq 1, \forall u \in \Sigma^+, \forall p, q \in Q, \quad p \xrightarrow{u^n} q \iff p \xrightarrow{u^{n+1}} q$$



$$3 \xrightarrow{a^1} 4$$

$$3 \xrightarrow{a^4} 4$$

$$3 \xrightarrow{a^6} 4$$

$$3 \xrightarrow{a^7} 4$$

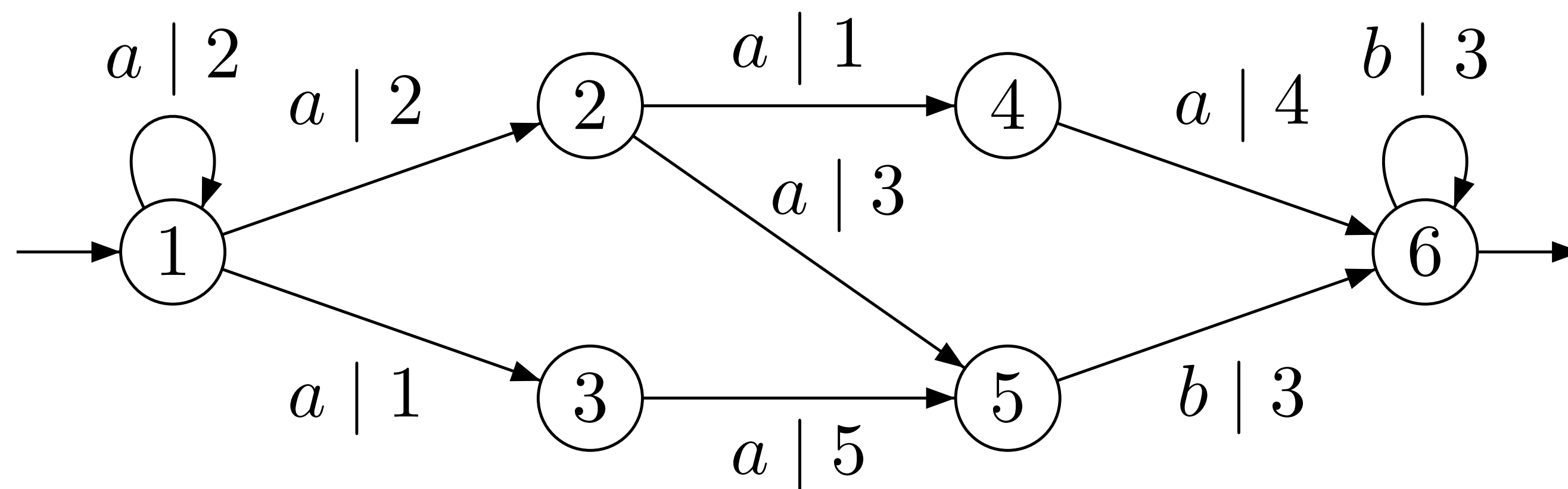
$$n = 6$$

Theorem (Boolean) [Schützenberger 65, McNaughton, Papert 71]

Aperiodic automata = FO sentences

Ambiguity in automata

k -ambiguous: Every word has **at most k accepting** run.



Domain: $a^*(a^3 + a^2b)b^*$
3-ambiguous

Input word: $w = a^n a^3 b b^p$

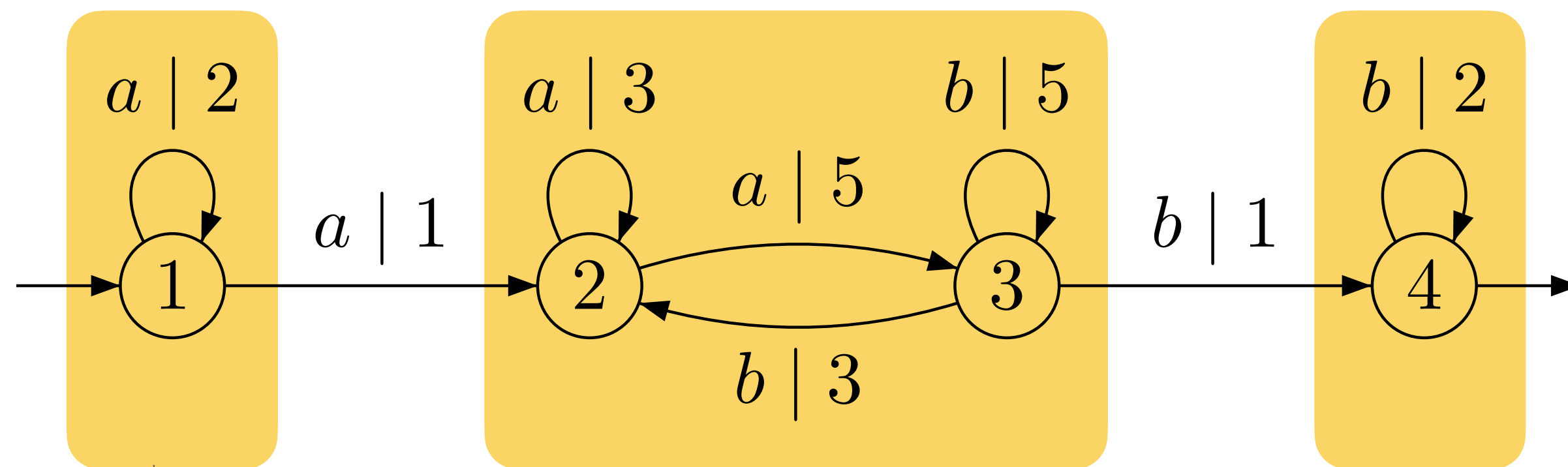
Unambiguous = 1-ambiguous

- $1 \xrightarrow{a} 1 \dots 1 \xrightarrow{a} 2 \xrightarrow{a} 4 \xrightarrow{a} 6 \xrightarrow{b} 6 \dots 6 \xrightarrow{b} 6$
- $1 \xrightarrow{a} 1 \dots 1 \xrightarrow{a} 1 \xrightarrow{a} 2 \xrightarrow{a} 5 \xrightarrow{b} 6 \dots 6 \xrightarrow{b} 6$
- $1 \xrightarrow{a} 1 \dots 1 \xrightarrow{a} 1 \xrightarrow{a} 3 \xrightarrow{a} 5 \xrightarrow{b} 6 \dots 6 \xrightarrow{b} 6$

Ambiguity in automata

Polynomially ambiguous:

Every word w has at most $\mathcal{O}(|w|^k)$ accepting run.



Domain: $a^+(a^+b^+)^+$
Quadratic ambiguity

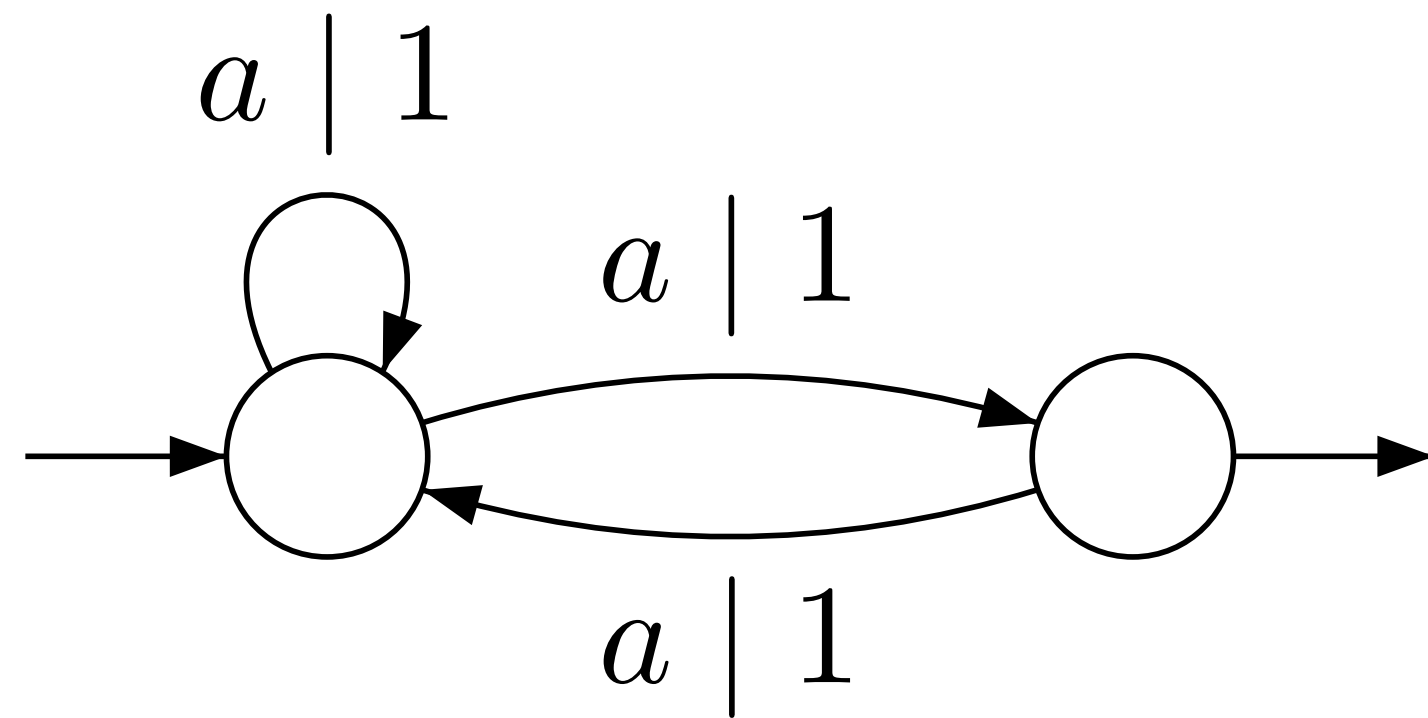
Polynomially ambiguous
iff
unambiguous on SCCs

Input word: $w \in a^n a(b^+a^+)^* b^p$

- $n \times p$ accepting runs on w
- Less than $|w|^2$ accepting runs on w

Ambiguity in automata

General case: Exponential ambiguity



$\mathcal{A}$ is aperiodic: index 2

Domain: a^+

F_n : number of accepting runs on a^n

$$F_0 = 0, F_1 = 1, F_{n+2} = F_{n+1} + F_n$$

Fibonacci numbers: exponential ambiguity

Aperiodic weighted Automata

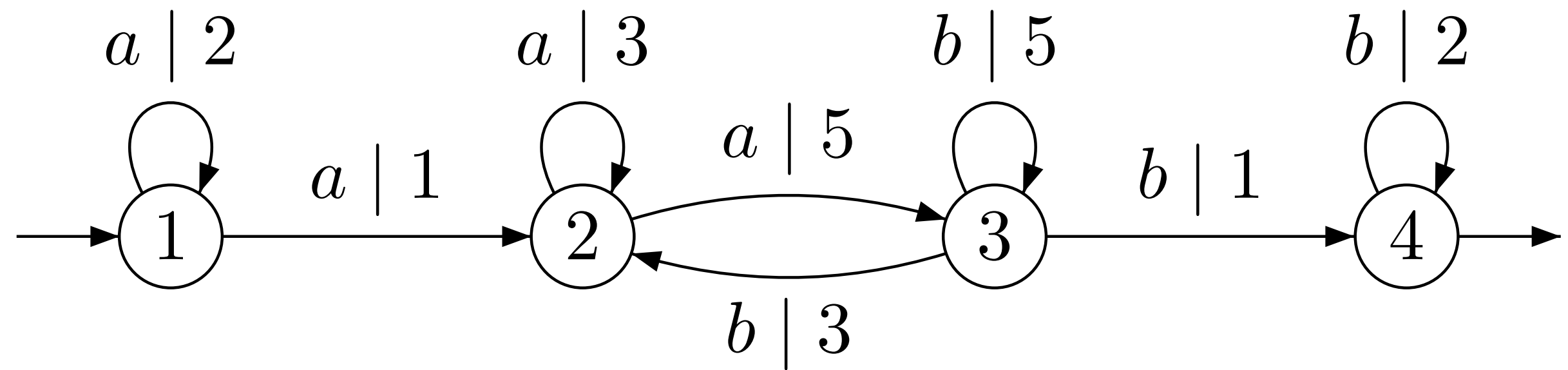
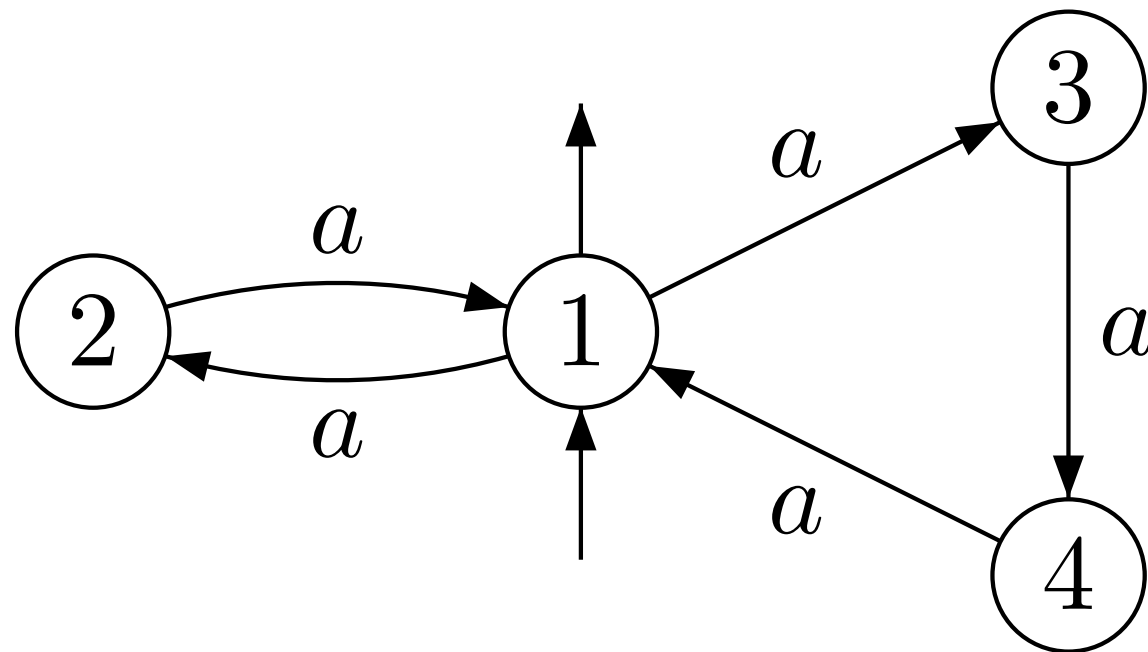
Manfred Droste, P.G., 2019

Strict ↑	Exponentially ambiguous		
	Polynomially ambiguous	$\Leftrightarrow$	wFO
	Finitely ambiguous	$\Leftrightarrow$	wFO without $\sum_x$
	Unambiguous	$\Leftrightarrow$	wFO without $\sum_x$ and without +

Aperiodic vs Counter-free

Aperiodic: $\exists n \geq 1, \forall u \in \Sigma^+, \forall p, q \in Q, p \xrightarrow{u^n} q \iff p \xrightarrow{u^{n+1}} q$

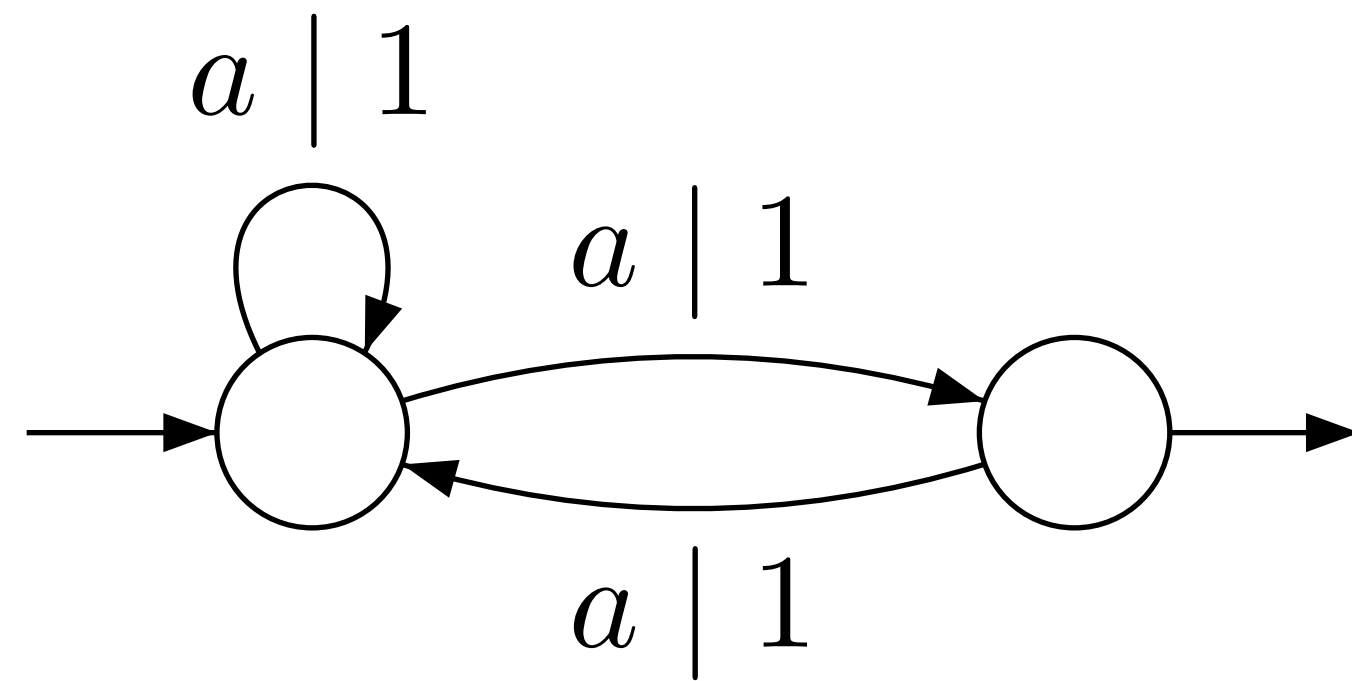
Counter-free: $\forall k \geq 1, \forall u \in \Sigma^+, \forall p \in Q, p \xrightarrow{u^k} p \implies p \xrightarrow{u} p$



Counter-free $\implies$ Aperiodic

Aperiodic $\implies$ Counter-free if $\mathcal{A}$ is polynomially ambiguous (SCC-unambiguous)

Poly-ambiguous vs Exp-ambiguous



Fibonacci numbers: $[[\mathcal{A}]](a^n) = F_n$

Assume $[[\mathcal{B}]](a^n) = F_n$ for some aperiodic polynomially ambiguous wA.

Polynomially ambiguous & Aperiodic $\implies$ SCC-unambiguous & counter-free

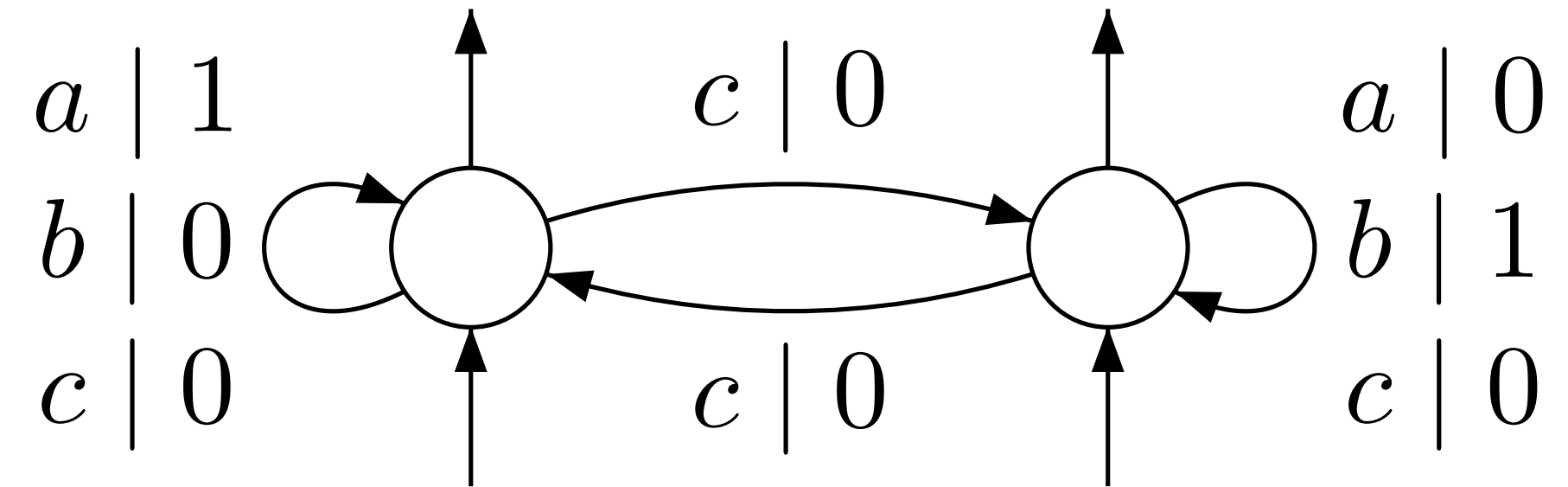
SCC-unambiguous & counter-free & $\Sigma = \{a\} \implies$ SCCs are just self-loops

$\mathbb{N}_{+, \times} = (\mathbb{N}, +, \times, 0, 1)$ semiring & $F_n = o(2^n) \implies \text{weight}(\text{self-loop}) \leq 1$

Then, $\text{wt}(\rho) \leq K$ for some constant $K \in \mathbb{N}$

Therefore, $[[\mathcal{B}]](a^n) \leq K \times \text{poly}(n) = o(F_n)$, a contradiction.

Poly-ambiguous vs Exp-ambiguous



$\mathbb{N}_{\max,+} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$: max-plus semiring.

$\mathcal{A}$ is exponentially ambiguous & counter-free.

Let $w = w_0 c w_1 c \cdots c w_n$ with $w_i \in \{a, b\}^*$. Then,

$$\llbracket \mathcal{A} \rrbracket(w) = \sum_{i=0}^n \max\{|w_i|_a, |w_i|_b\}.$$

This function cannot be computed with an aperiodic poly-ambiguous wA.

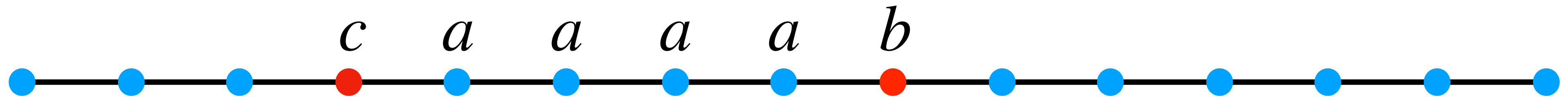
Aperiodic weighted Automata

Manfred Droste, P.G., 2019

Strict ↑	Exponentially ambiguous		
	Polynomially ambiguous	$\Leftrightarrow$	wFO
	Finitely ambiguous	$\Leftrightarrow$	wFO without $\sum_x$
	Unambiguous	$\Leftrightarrow$	wFO without $\sum_x$ and without +

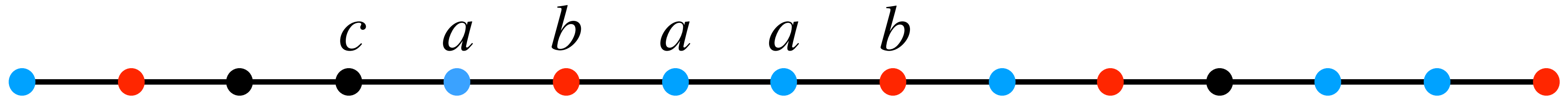
Linear Temporal Logic (LTL)

$\varphi ::= \top \mid a \mid \neg\varphi \mid \varphi \vee \varphi \mid X\varphi \mid \varphi \cup \varphi \mid Y\varphi \mid \varphi S \varphi$



Linear Temporal Logic (LTL)

$$\varphi ::= \top \mid a \mid \neg\varphi \mid \varphi \vee \varphi \mid X\varphi \mid \varphi \cup \varphi \mid Y\varphi \mid \varphi S \varphi$$



Example: $L = (ab + aab + c)^+$ is LTL-definable

- $F\varphi = \top \cup \varphi$ and $G\varphi = \neg F\neg\varphi$

Theorem [Kamp, PhD 1968]

FO = LTL.

Weighted Linear Temporal Logic (wLTL)

$$\text{LTL} \quad \varphi ::= \top \mid a \mid \neg \varphi \mid \varphi \vee \varphi \mid X \varphi \mid \varphi \cup \varphi \mid Y \varphi \mid \varphi S \varphi$$

$$\text{step-wLTL} \quad \alpha ::= r \mid \varphi ? \alpha : \alpha$$

$$\text{wLTL} \quad \Phi ::= \mathbf{0} \mid \mathbf{L} \alpha \mid \langle \alpha \rangle \Phi \mid \alpha \cup \Phi \mid \Phi + \Phi \mid \varphi ? \Phi : \Phi$$

Semantics: $w \in \Sigma^+$ and $i \in \text{pos}(w)$

$$\llbracket r \rrbracket_i^w = r$$

$$\llbracket \varphi ? \alpha_1 : \alpha_2 \rrbracket_i^w = \begin{cases} \llbracket \alpha_1 \rrbracket_i^w & \text{if } w, i \models \varphi \\ \llbracket \alpha_2 \rrbracket_i^w & \text{otherwise.} \end{cases}$$

$$\llbracket \mathbf{L} \alpha \rrbracket_i^w = \begin{cases} \llbracket \alpha \rrbracket_i^w & \text{if } i = |w| \\ 0 & \text{otherwise.} \end{cases}$$

$$\llbracket \langle \alpha \rangle \Phi \rrbracket_i^w = \begin{cases} \llbracket \alpha \rrbracket_i^w \cdot \llbracket \Phi \rrbracket_{i+1}^w & \text{if } i < |w| \\ 0 & \text{otherwise.} \end{cases}$$

$$\llbracket \alpha \cup \Phi \rrbracket_i^w = \sum_{i \leq k \leq |w|} \left(\prod_{i \leq j < k} \llbracket \alpha \rrbracket_j^w \right) \times \llbracket \Phi \rrbracket_k^w$$

Weighted Linear Temporal Logic (wLTL)

LTL $\varphi ::= \top \mid a \mid \neg\varphi \mid \varphi \vee \varphi \mid X\varphi \mid \varphi \cup \varphi \mid Y\varphi \mid \varphi S \varphi$

step-wLTL $\alpha ::= r \mid \varphi ? \alpha : \alpha$

wLTL $\Phi ::= \mathbf{0} \mid \mathbf{L} \alpha \mid \langle \alpha \rangle \Phi \mid \alpha \cup \Phi \mid \Phi + \Phi \mid \varphi ? \Phi : \Phi$

Example: $L = (ab + aab + c)^+$ is LTL-definable

- $\varphi_L = \mathbf{G}((b \rightarrow Y a) \wedge (a \rightarrow (X b \vee X(a \wedge X b))))$

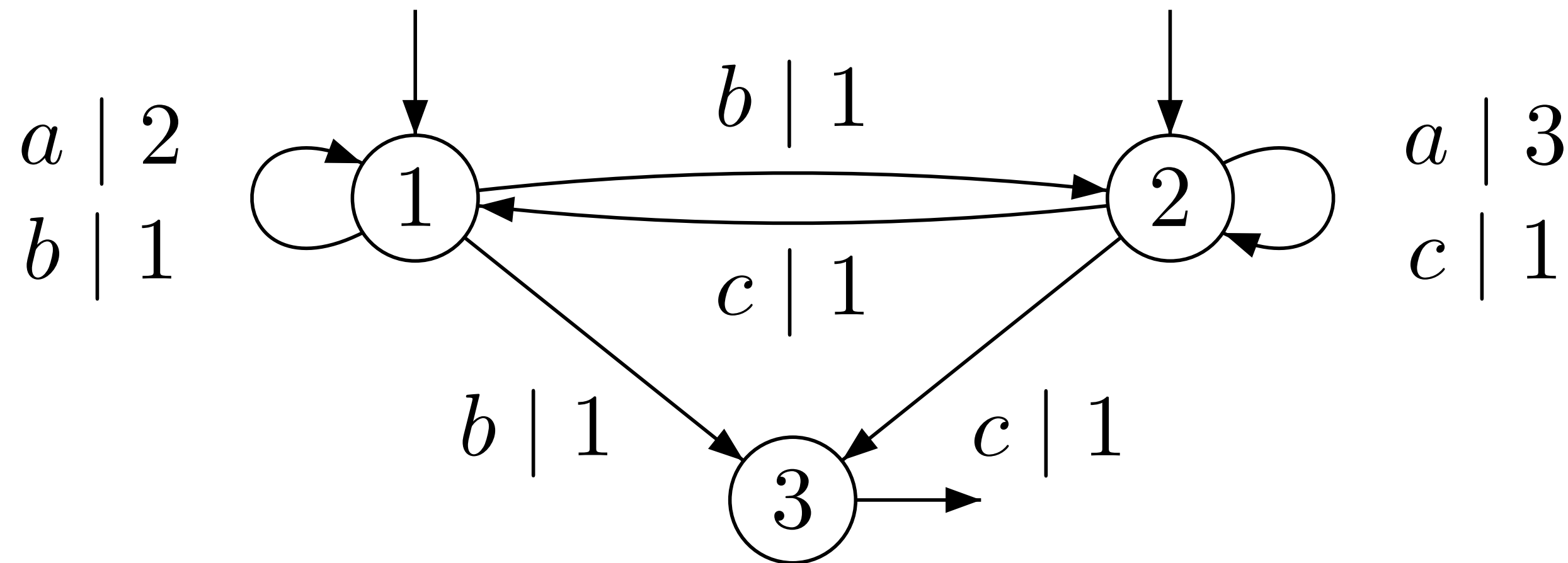
- $\alpha = (a \vee c) ? 1 : ((Y Y a) ? 3 : 2)$

- $\Phi = \varphi_L ? \mathbf{G} \alpha : \mathbf{0}$

$\mathbf{G} \alpha = \alpha \cup \mathbf{L} \alpha$

- For $w \in L$, let $n = |w|_{ab}$ and $m = |w|_{aab}$. Then, $[\![\Phi]\!]_1^w = 3^m \times 2^{n-m}$

Unambiguous aperiodic wA vs wLTL



Domain of $\mathcal{A}$: $L = (a^*b + a^*c)^+ = \Sigma^*(b + c)$

- $\varphi_L = \mathbf{G}(b \vee c \vee \mathbf{X} \top)$
- $\alpha = (b \vee c) ? 1 : ((a \cup b) ? 2 : 3)$
- $\Phi = \varphi_L ? \mathbf{G} \alpha : \mathbf{0}$
- Then, $\llbracket \Phi \rrbracket_1^w = \llbracket \mathcal{A} \rrbracket(w)$ for all $w \in \Sigma^+$

wLTL(G) without + and U

LTL $\varphi ::= \top \mid a \mid \neg\varphi \mid \varphi \vee \varphi \mid X\varphi \mid \varphi \cup \varphi \mid Y\varphi \mid \varphi S \varphi$

step-wLTL $\alpha ::= r \mid \varphi ? \alpha : \alpha$

wLTL $\Phi ::= \mathbf{0} \mid \mathbf{G} \alpha \mid \langle \alpha \rangle \Phi \mid \varphi ? \Phi : \Phi$

$\mathbf{G} \alpha = \alpha \cup \mathbf{L} \alpha$

Semantics: $w \in \Sigma^+$ and $i \in \text{pos}(w)$ $\llbracket \mathbf{G} \alpha \rrbracket_i^w = \prod_{i \leq j \leq |w|} \llbracket \alpha \rrbracket_j^w$

wLTL(G) without + and U is contained in wFO
without binary (+) or first-order ($\sum_x$) sums.

Theorem [Kamp, PhD 1968]

FO = LTL.

For each α in step-wLTL there is an equivalent $\alpha'(z)$ in step wFO: $\llbracket \alpha \rrbracket_i^w = \llbracket \alpha' \rrbracket_{[z \mapsto i]}^w$

We get $\llbracket \mathbf{G} \alpha \rrbracket_1^w = \prod_{1 \leq i \leq |w|} \llbracket \alpha \rrbracket_i^w = \prod_{1 \leq i \leq |w|} \llbracket \alpha' \rrbracket_{[z \mapsto i]}^w = \llbracket \prod_z \alpha' \rrbracket^w$

Unambiguous aperiodic wA vs wLTL

Let $\mathcal{A}$ be an aperiodic and unambiguous wA

- $L = \text{dom}(\mathcal{A})$ is LTL-definable by some sentence φ_L in LTL
- For states p, q we construct $\psi_p \in \text{LTL}(Y, S)$ and $\phi_q \in \text{LTL}(X, U)$ such that

$$I \xrightarrow{u} p \text{ iff } u, |u| \models \psi_p \quad \text{and} \quad q \xrightarrow{v} F \text{ iff } v, 1 \models \phi_q$$

- For each transition $\delta = p \xrightarrow{a} q$, let $\varphi_\delta = Y \psi_p \wedge a \wedge X \phi_q$.

For all $w \in L$ and $i \in \text{pos}(w)$

$$w, i \models \varphi_\delta \text{ iff } \delta \text{ is the } i\text{-th transition in the accepting run of } \mathcal{A} \text{ on } w$$

- Construct α in step-wLTL such that $\llbracket \alpha \rrbracket_i^w = \text{wt}(\delta)$ when $w, i \models \varphi_\delta$
- Let $\Phi = \varphi_L ? G \alpha : \mathbf{0}$. Then, $\llbracket \Phi \rrbracket_1^w = \llbracket \mathcal{A} \rrbracket(w)$ for all $w \in \Sigma^+$

Unambiguous aperiodic wA vs wLTL

Theorem: The following classes are expressively equivalent:

1. Aperiodic unambiguous wA
2. wFO sentences without binary (+) or first-order ($\sum_x$) sums.
3. wLTL(G) sentences without binary (+) sums or weighted until.

We just saw $(1) \implies (3)$ and $(3) \implies (2)$.

$(2) \implies (1)$ proved in Droste-Gastin 2019.

Aperiodic wA vs wFO vs wLTL

Exponentially ambiguous			
Polynomially ambiguous	$\longleftrightarrow$	wFO	$\overset{?}{\longleftrightarrow}$ wLTL
Finitely ambiguous	$\longleftrightarrow$	wFO without $\sum_x$	$\longleftrightarrow$ wLTL without U
Unambiguous	$\longleftrightarrow$	wFO without + and $\sum_x$	$\longleftrightarrow$ wLTL no + and U

wFO vs wLTL

Example: Let $L \subseteq \Sigma^+$ be a language defined by the FO sentence φ .

$$\Phi_{\text{pref}} = \sum_x \varphi^{\leq x} ? \mathbf{1} : \mathbf{0}$$

where $\mathbf{1} = \prod_z 1$ and $\varphi^{\leq x}$ is φ relativized to the prefix $[1, x]$

Number of
prefixes in L

Example: Let $L \subseteq \Sigma^+$ be a language defined by the formula $\psi \in \text{LTL}(Y, S)$.

$$\Phi = (1 \cup \langle \psi ? 1 : \mathbf{0} \rangle G 1) + (1 \cup L(\psi ? 1 : \mathbf{0}))$$

wFO vs wLTL

Example: Let $L \subseteq \Sigma^+$ be a language defined by the FO sentence φ .

$$\Phi_{\text{inf}} = \sum_{x,y} (x \leq y \wedge \varphi^{[x,y]}) ? \mathbf{1} : \mathbf{0}$$

where $\mathbf{1} = \prod_z 1$ and $\varphi^{[x,y]}$ is φ relativized to the infix $[x, y]$

Number of
factors in L

Thank you for your attention!