Star Height of Reversible Languages and Universal Automata

Sylvain Lombardy and Jacques Sakarovitch

CNRS / ENST
Paris - France

Star height

Star height of an expression:
Number of nested calls to the star operator.

$$\mathsf{h}(a + a^2(b + a^3)^* a^2) = 1, \qquad \mathsf{h}(a(ab^*a^2)^*) = 2$$

Star height of a regular language $\mathcal{L}$:
Minimal star height of expressions that denote $\mathcal{L}$.

$$\mathcal{L} = |a + a^2(b + a^3)^* a^2| = |a(ab^*a^2)^*| \longrightarrow \mathsf{h}(\mathcal{L}) = 1$$

$\rightarrow$ **Question** (Eggan 63):
Is star height of regular languages computable ?

Partial solutions

– *Pure group languages*: McNaughton, 67
– *Reset free languages*: Hashiguchi & Honda, 79
– *Strictly locally testable languages*: H & H, 79

– Star height *1* decidable: Hashiguchi, 82

Solution

Th (Hashiguchi, 88): Star height is *computable*.

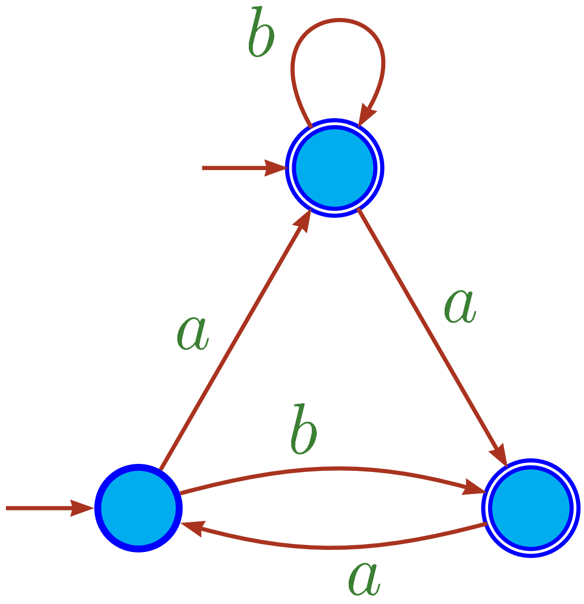
$\rightarrow$ *Enumeration* of regular expressions with bounded length: no *structural* information

$\rightarrow$ *Untractable* algorithm:
Number of regular expressions to be checked for a language that will be computed of star height 2:

$$N > 2^{2^{2^{2^{2^{2^{2^{2^{2}}}}}}}} > 10^{10^{10^{10^{10^{10}}}}}$$

Reversible automaton

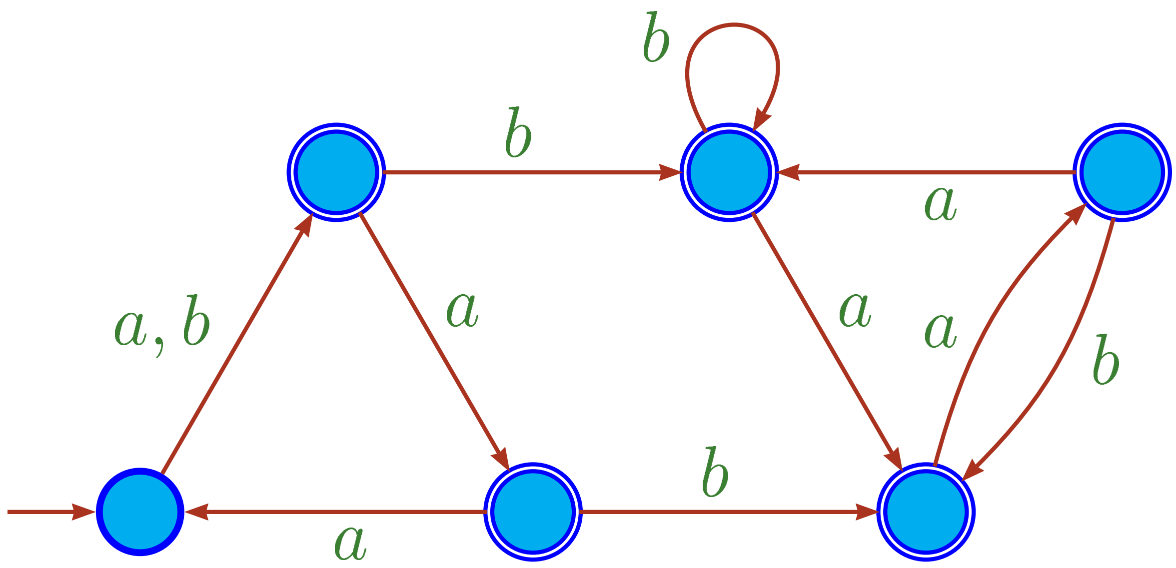
– Transition deterministic &
codeterministic automaton
– Several initial and terminal states



Reversible language

$\mathcal{L}$ reversible $\Leftrightarrow \exists\ \mathcal{A}$ reversible, $L(\mathcal{A}) = \mathcal{L}$.

Pb: $\mathcal{L}$ rev. $\not\Rightarrow$ minimal automaton of $\mathcal{A}$ is reversible



Th (Pin, Latin'92): Reversibility is decidable in polynomial time ($O(n^4)$).

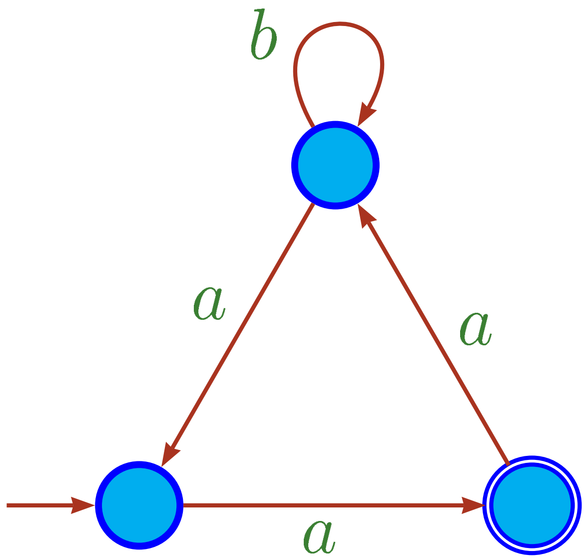
Star height of reversible language

Th 1: The star height of *reversible languages* is computable (with a tractable algorithm).

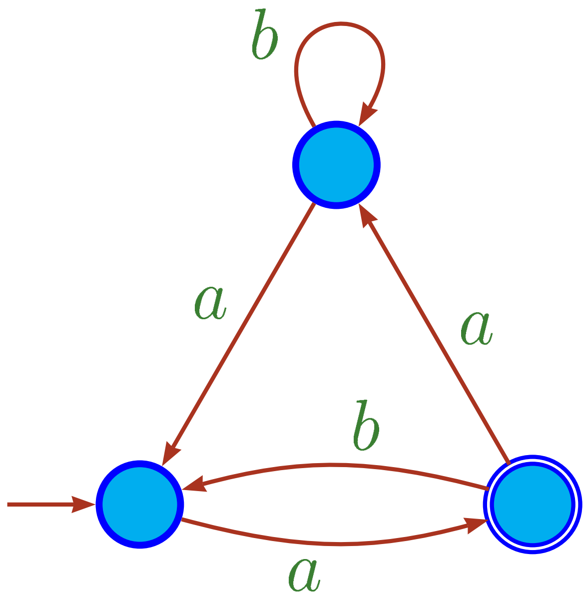
Pure group and *reset-free* languages are special cases of *reversible* languages.

Loop complexity

= Number of “nested circuits”

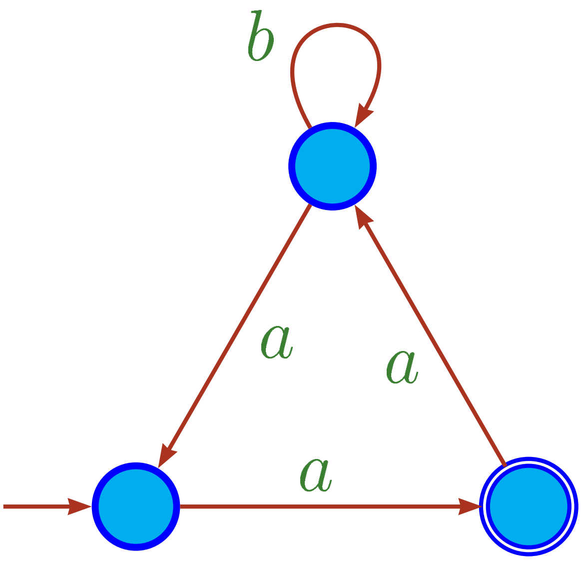


lc $= 1$



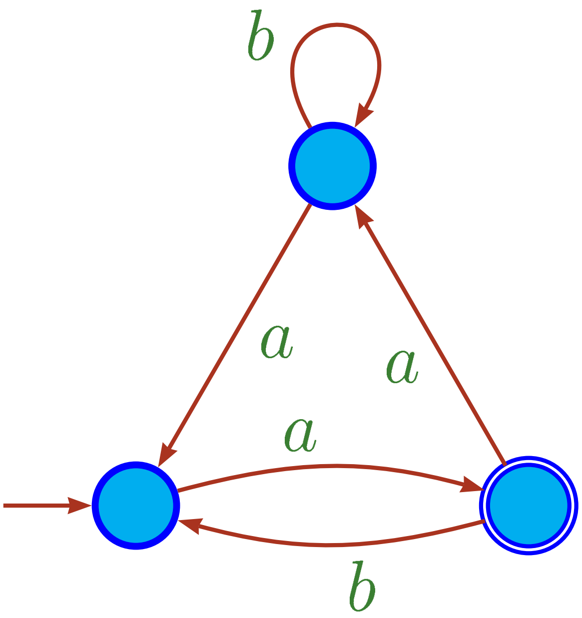
lc $= 2$

Star height $\longleftrightarrow$ Loop complexity



$\rightarrow \mathsf{E} = a + a(b + aaa)^* a$

$\mathsf{h}(\mathsf{E}) = 1$



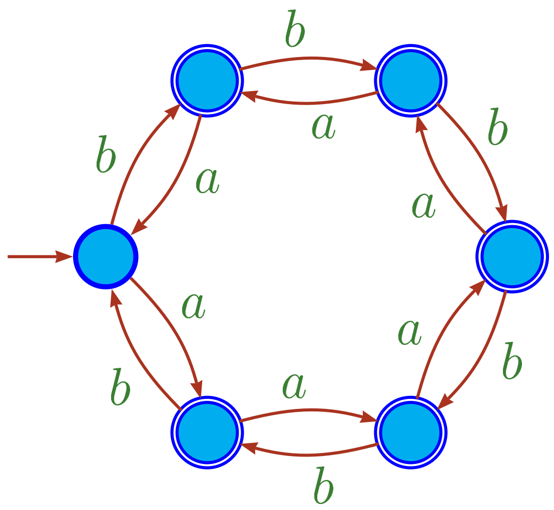
$\rightarrow \mathsf{E} = (a(b + ab^* a))^* a$

$\mathsf{h}(\mathsf{E}) = 2$

Th (Eggan, 63): $\mathsf{h}(\mathcal{L}) = \min\{\mathsf{lc}(\mathcal{A}) \mid L(\mathcal{A}) = \mathcal{L}\}$

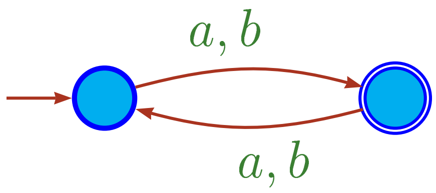
Pb: The loop complexity of the minimal automaton of $\mathcal{L}$ may be *not* minimal.

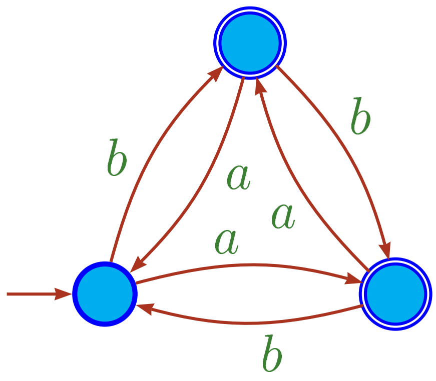
$$\mathcal{L} = \{u \mid |u|_a \neq |u|_b \mod 6\}$$



$\longrightarrow \mathsf{lc} = 3$

$$|u|_a \neq |u|_b \mod 6 \Leftrightarrow \begin{cases} \text{or} & |u|_a \neq |u|_b \mod 2 \\ & |u|_a \neq |u|_b \mod 3 \end{cases}$$





$\longrightarrow \mathsf{lc} = 2$

$\longrightarrow \mathsf{h}(\mathcal{L}) = 2$

Star height of reversible language

Th 1: The star height of *reversible languages* is computable (with a tractable algorithm).

Pure group and *reset-free* languages are special cases of *reversible* languages.

Extension of proofs of
– McNaughton (67),
– Hashiguchi & Honda (79),
by means of the *universal automaton*.

Universal automaton

Conway, 1971: *Regular algebra and finite machines*
$\rightarrow$ "Factor matrix"

For every regular language $\mathcal{L}$,

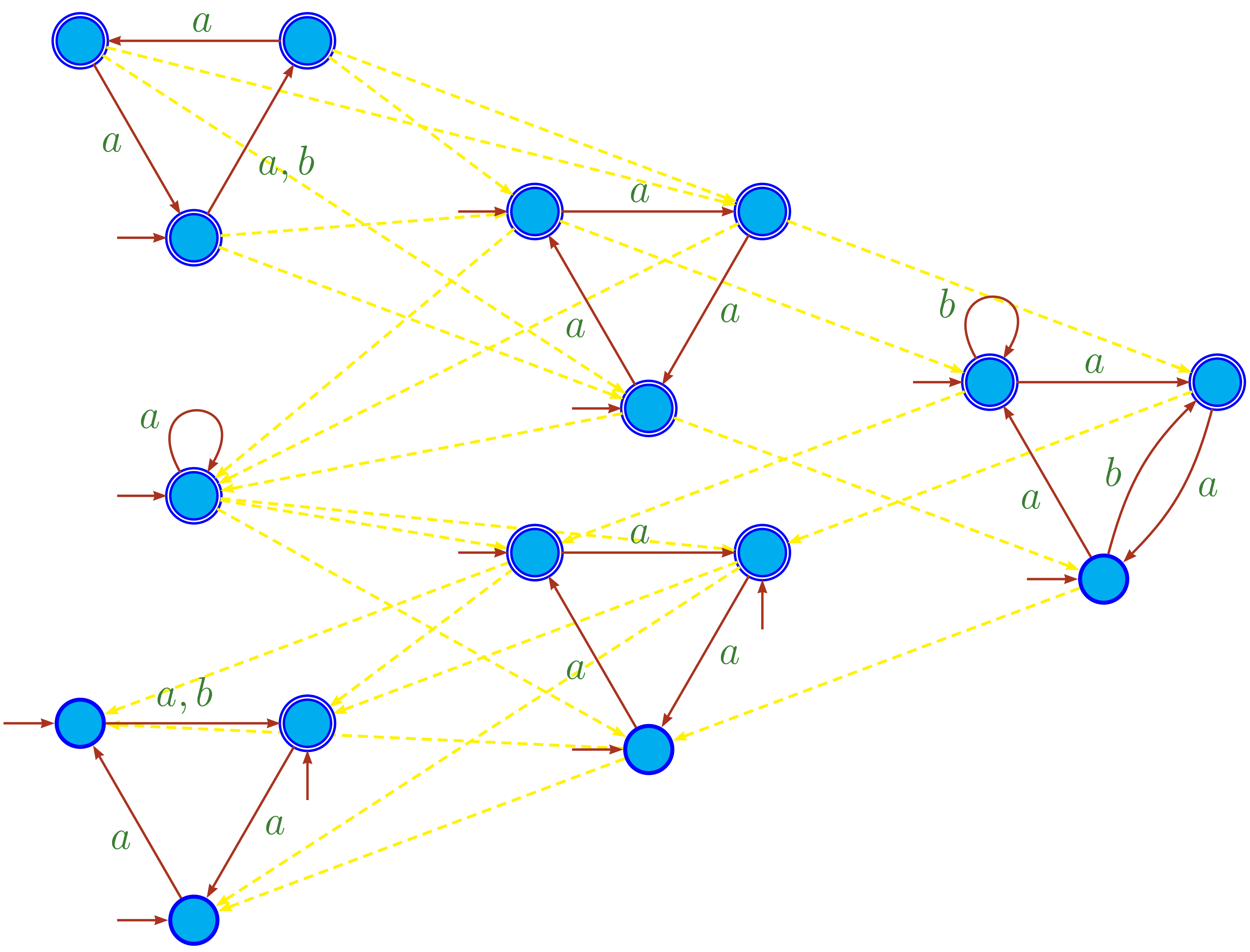
there exists a finite automaton $\mathcal{U}_{\mathcal{L}}$ such that :

– $\mathcal{U}_{\mathcal{L}}$ accepts $\mathcal{L}$;
– For every DFA or NFA $\mathcal{A}$ that accepts $\mathcal{L}$,
there exists a morphism $\varphi : \mathcal{A} \longrightarrow \mathcal{U}_{\mathcal{L}}$;
– $\mathcal{U}_{\mathcal{L}}$ is minimal w.r.t. this property.

Computing star height with the universal automaton

Theorem 2
The universal automaton of a reversible language contains an *equivalent subautomaton* with *minimal loop complexity*.

– $\mathcal{U}_{\mathcal{L}}$ is computable,
– the size of $\mathcal{U}_{\mathcal{L}}$ is exponential.

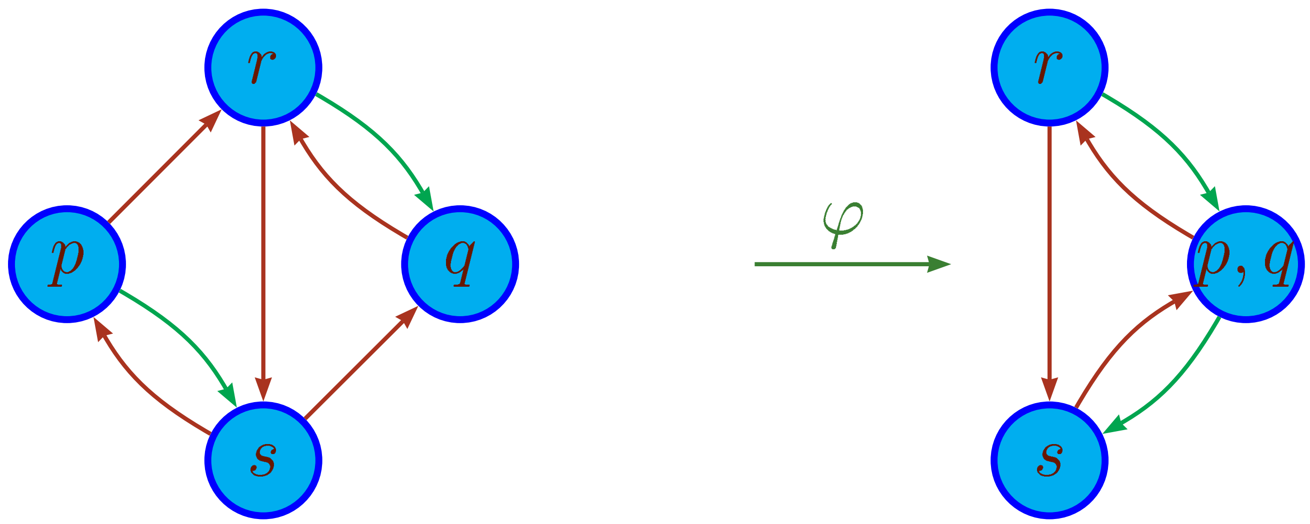


a
a
a, b
a
a
a
b
a
a
b
a
a
a
a
a
a, b
a
a

Steps of proof

Prop: Strongly connected components of the universal automaton of a reversible language are reversible.

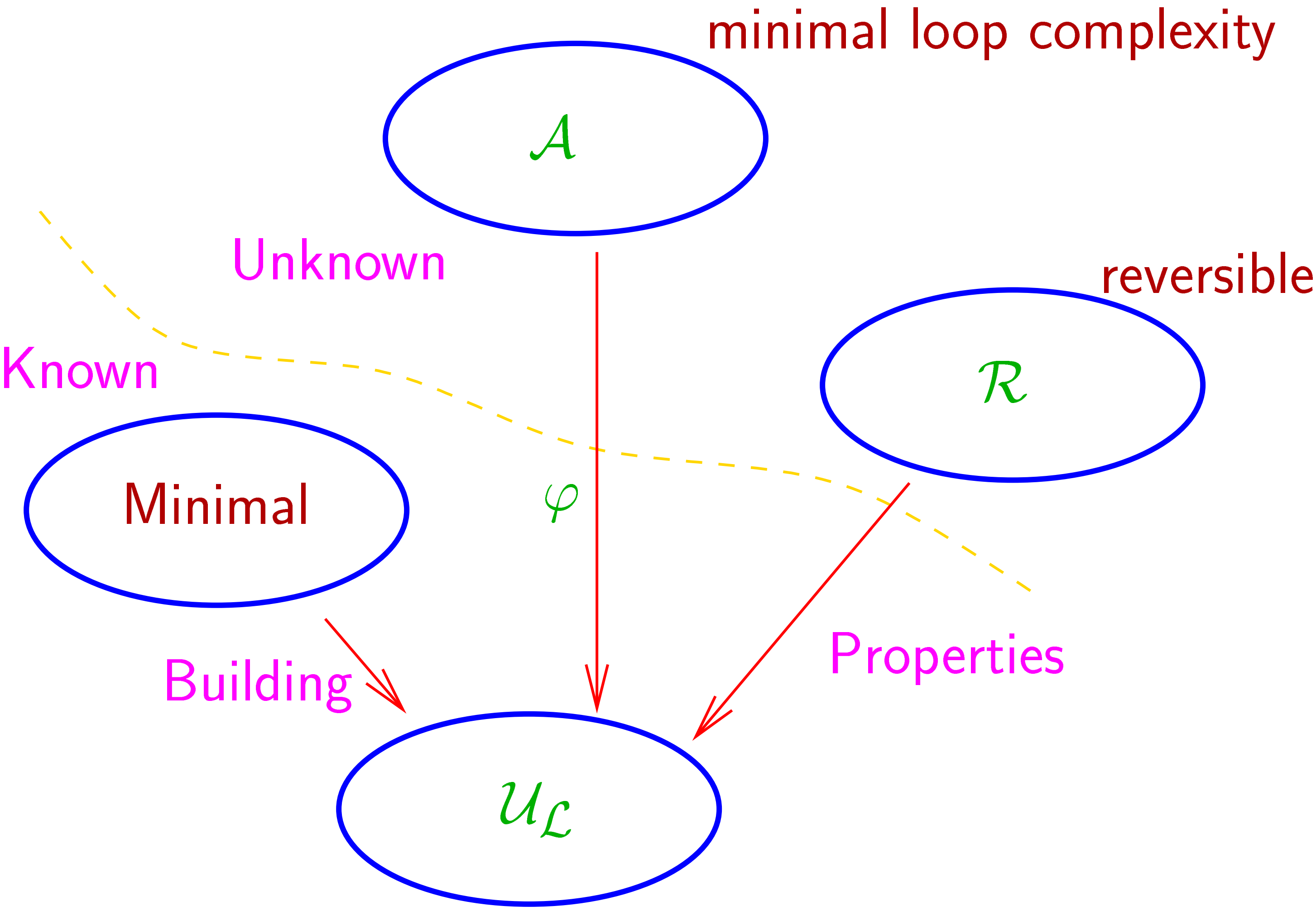
Def: Conformal morphisms.



Non conformal morphism

Prop (McNaughton, 67):
$\varphi : \mathcal{A} \to \mathcal{B}$ conformal $\Rightarrow \mathsf{lc}(\mathcal{B}) \leqslant \mathsf{lc}(\mathcal{A})$.

Steps of proof (2)



Steps of proof (3)

$\mathcal{L}$ reversible language

– $\exists\ \mathcal{R}$ reversible automaton (unknown)
– $\exists\ \mathcal{A}$ automaton with minimal loop complexity (unknown)
– $\exists\ \mathcal{U}_{\mathcal{L}}$ universal automaton (well known !)
– $\exists\ \varphi : \mathcal{A} \to \mathcal{U}_{\mathcal{L}}$

$\mathcal{R} \to$ finite decomposition of $\mathcal{L} = \bigcup \mathcal{L}_i$

For every $\mathcal{L}_i$, $\exists \mathcal{A}_i$ subautomaton of $\mathcal{A}$ s.t.
– $\varphi_{|\mathcal{A}_i}$ is conformal $\Rightarrow \mathsf{lc}(\varphi(\mathcal{A}_i)) \leqslant \mathsf{lc}(\mathcal{A})$
– $\varphi(\mathcal{A}_i)$ accepts at least $\mathcal{L}_i$

We check that $\mathsf{lc}(\bigcup \varphi(\mathcal{A}_i)) \leqslant \mathsf{lc}(\mathcal{A})$
$\to$ This is a subautomaton of $\mathcal{U}_{\mathcal{L}}$ with minimal loop complexity !