Introduction

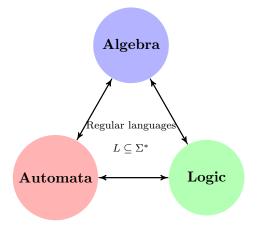
Automata, Logic and Algebra for (Finite) Word Transductions

Emmanuel Filiot

Université libre de Bruxelles & FNRS



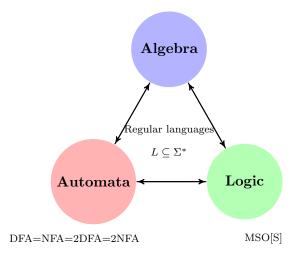
Trinity for Regular Languages

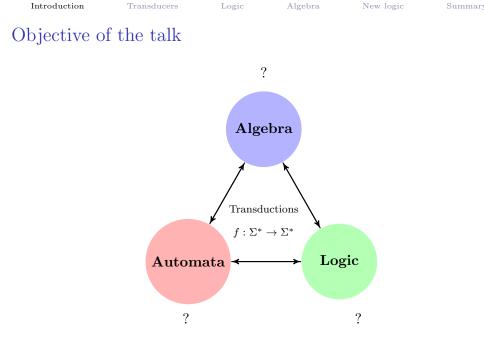


Introduction Transducers Logic Algebra New logic Summar

Trinity for Regular Languages

Finite monoids

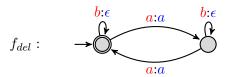




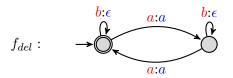
Introduction

Automata models for transductions

Introduction	Transducers	Logic	Algebra	New logic	Summary

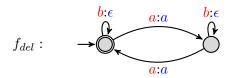


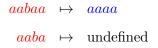
Introduction	Transducers	Logic	Algebra	New logic	Summary



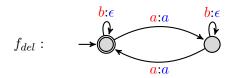
 $aabaa \mapsto aaaa$

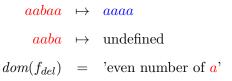
Introduction	Transducers	Logic	Algebra	New logic	Summary





Introduction	Transducers	Logic	Algebra	New logic	Summary



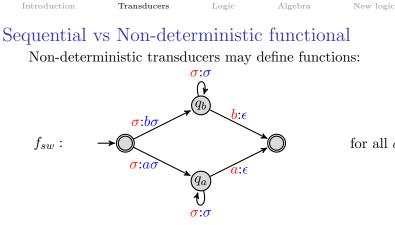


Introduction Transducers Logic Algebra New logic Summary Non-determinism

In general, transducers define binary relations in $\Sigma^* \times \Sigma^*$

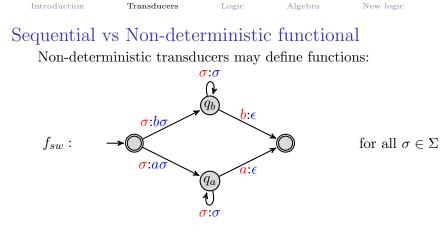


realizes $\{(u, v) \mid v \text{ is a subword of } u\}$



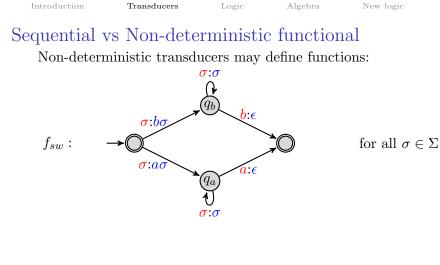


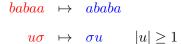
Summar



 $babaa \mapsto ababa$

Summar



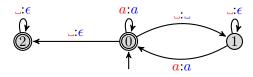


input-determinism (aka sequential) < non-determinism \cap functions

Summar

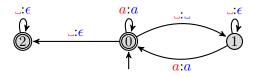
Introduction Transducers

 $_{-}$ = white space



Introduction Transducers	Logic	Algebra
--------------------------	-------	---------

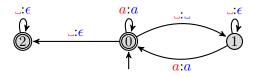
 $_=$ white space





Introduction T	ransducers Logi	c Algebra
----------------	-----------------	-----------

 $_{-}$ = white space

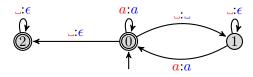




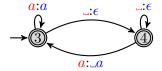
Is non-determinism needed ?

Introduction Transducers	Logic	Algebra	
--------------------------	-------	---------	--

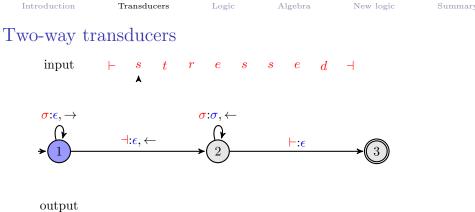
 $_{-}$ = white space



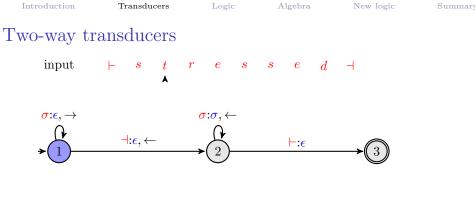


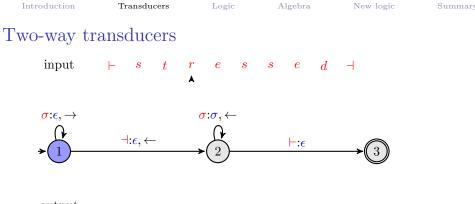


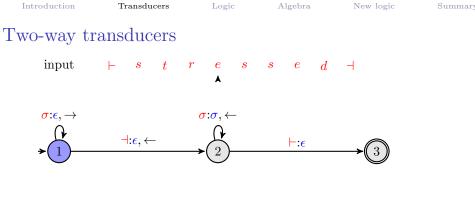
Is non-determinism needed ? No.

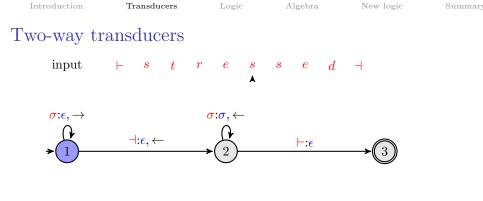


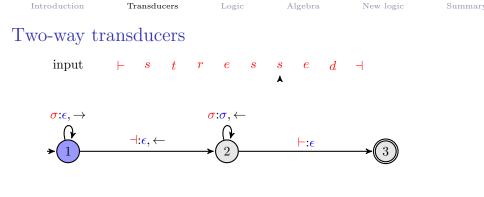
.

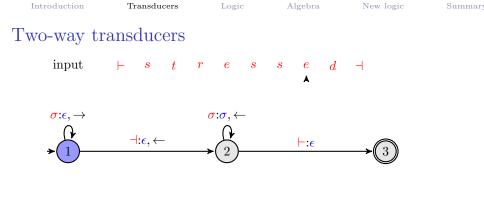




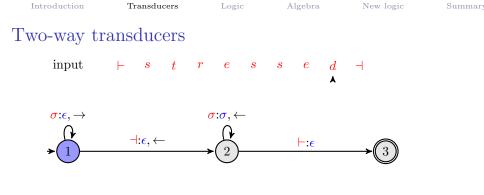




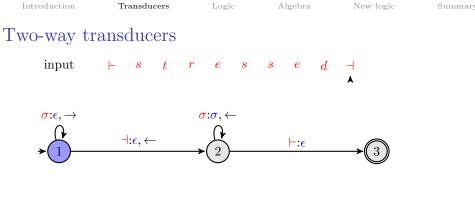


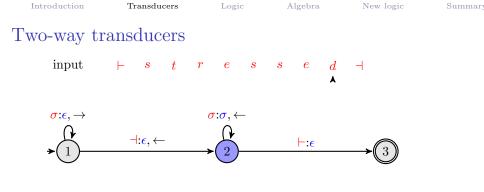


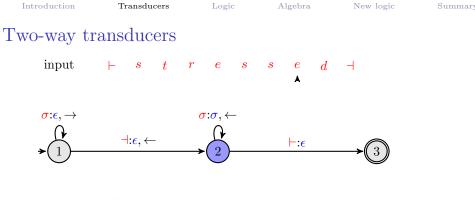
output



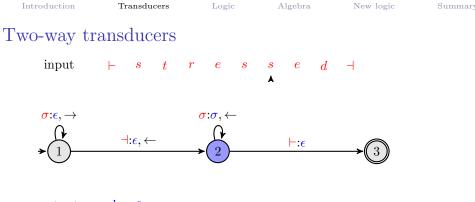
output



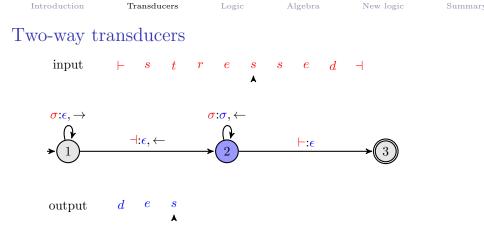


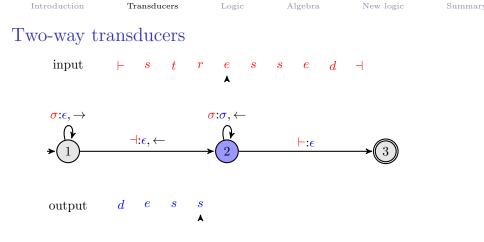


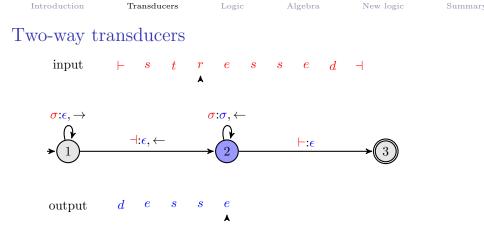
output d

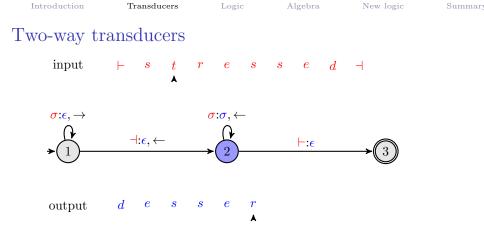


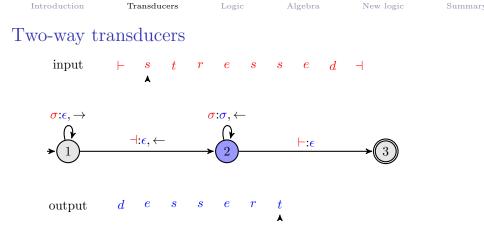
output d e

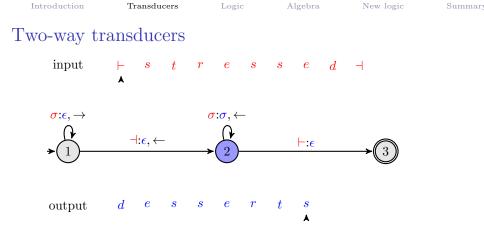


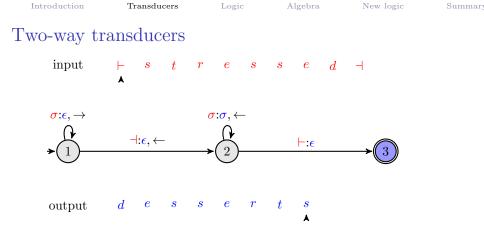


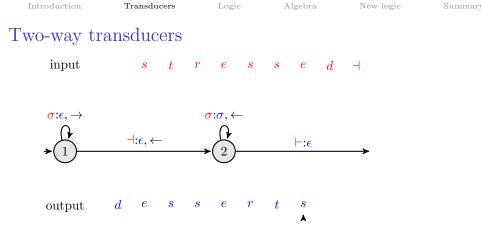










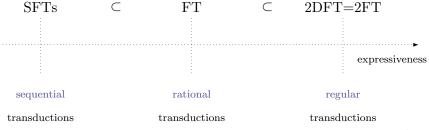


one-way < two-way

c) decidable equivalence problem (Culik, Karhumaki, 87).
c) closed under composition o (Chytil, Jakl, 77)

Introduction	Transducers	Logic	Algebra	New logic	Summary

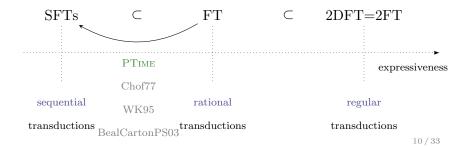
Landscape of Transducer Classes



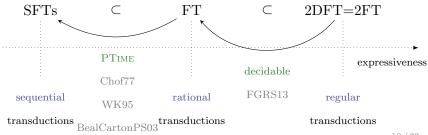
10/33

Introduction	Transducers	Logic	Algebra	New logic	Summary

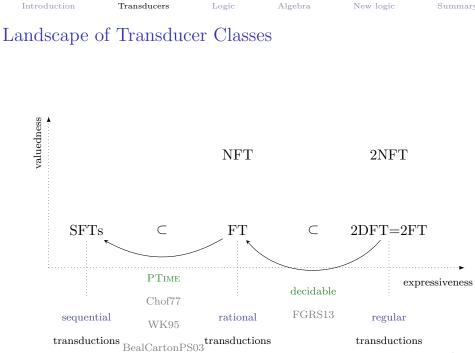
Landscape of Transducer Classes



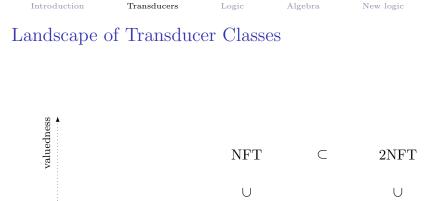
Landscape of Transducer Classes

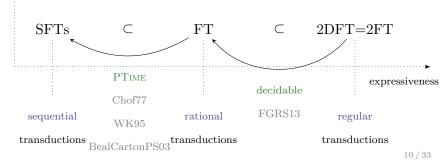


10/33



10/33





Summar

Introduction

Transducers

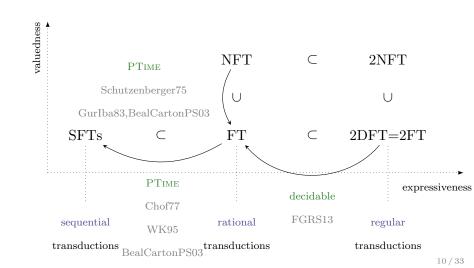
Logic

Algebra

New logic

Summar

Landscape of Transducer Classes



Introduction

Transducers

Logic

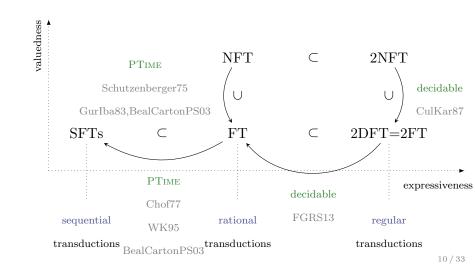
Algebra

a

New logic

Summar

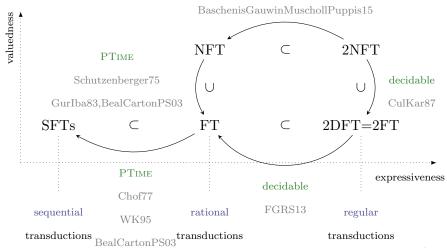
Landscape of Transducer Classes





undecidable

Landscape of Transducer Classes



10/33



- A surface back to a DET if some loss on datas (A
- ▶ equivalent to 2DFT if copyless updates (Alur, Cerny, 10)
- ▶ decidable equivalence problem (F., Reynier) ~ HDT0L



- deterministic one-way
- ▶ equivalent to 2DFT if copyless updates (Alur, Cerny, 10)
- ▶ decidable equivalence problem (F., Reynier) \sim HDT0L
- regular expressions to register transducer, implemented in DReX (Alur, D'Antoni, Raghothaman, 2015)
- register minimization for a subclass (Baschenis, Gauwin, Muscholl, Puppis, 16)



- deterministic one-way
- ▶ equivalent to 2DFT if copyless updates (Alur, Cerny, 10)
- ▶ decidable equivalence problem (F., Reynier) ~ HDT0L
- regular expressions to register transducer, implemented in DReX (Alur, D'Antoni, Raghothaman, 2015)
- register minimization for a subclass (Baschenis, Gauwin, Muscholl, Puppis, 16)

Two-way to one-way transducers

- ▶ decidable, but non-elementary complexity in (FGRS13)
- elementary complexity first obtained for subclasses (sweeping) by (BGMP15)
- ▶ recently for the full class (two talks at Delta)

Algebra

New logic

Summar

Logic for transductions

(Courcelle) MSO Transformations

"interpreting the output structure in the input structure"

(Courcelle) MSO Transformations

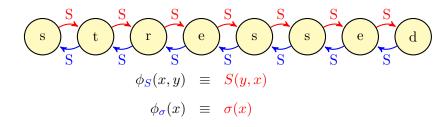
"interpreting the output structure in the input structure"

(Courcelle) MSO Transformations

"interpreting the output structure in the input structure"

(Courcelle) MSO Transformations

"interpreting the output structure in the input structure"



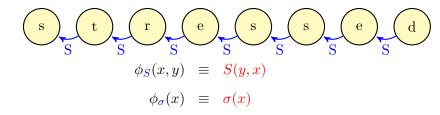
(Courcelle) MSO Transformations

"interpreting the output structure in the input structure"

(Courcelle) MSO Transformations

"interpreting the output structure in the input structure"

 output predicates defined by MSO[S] formulas interpreted over the input structure



• input structure can be copied a fixed number of times: $u \mapsto uu$, or $u \mapsto u$.mirror(u).

Introdu	ction	Transducers	Lo	gic Alge	bra New	v logic Summary
Büchi	Theorem	ns for W	Vord	Transduc	tions	

Let $f: \Sigma^* \to \Sigma^*$.

Theorem (Engelfriet, Hoogeboom, 01) f is 2FT-definable iff f is MSO-definable.

Introduction	Transducers	Logic	Algebra	New logic	Summary
Büchi Theo	rems for We	ord Tra	nsduction	S	

Let $f: \Sigma^* \to \Sigma^*$.

Theorem (Engelfriet, Hoogeboom, 01) f is 2FT-definable iff f is MSO-definable.

Consequence Equivalence is decidable for MSO-transducers.

Introduction	Transducers	Logic	Algebra	New logic	Summary
	C XX 7	1 00	1		
Büchi Theo	rems for We	ord Trai	isductions	5	

Let $f: \Sigma^* \to \Sigma^*$.

Theorem (Engelfriet, Hoogeboom, 01) f is 2FT-definable iff f is MSO-definable.

Consequence Equivalence is decidable for MSO-transducers.

Theorem (Bojanczyk 14, F. 15) f is (1)FT-definable iff f is order-preserving MSO-definable.

Order-preserving MSO: $\phi_S^{i,j}(x,y) \models x \preceq y$.

First-order transductions

Replace MSO by FO formulas.

Results

- ▶ equivalent to aperiodic transducers with registers (F., Trivedi, Krishna S., 14)
- ▶ and to aperiodic 2DFT (Carton, Dartois, 15) (Dartois, Jecker, Reynier, 16)

Introduction

Algebraic characterizations of transductions

- $u \sim_L v$ if: for all $w \in \Sigma^*$, $uw \in L$ iff $vw \in L$
- $\blacktriangleright\ u$ and v have the same "effect" on continuations w
- Myhill-Nerode's Thm: L is regular iff $\Sigma^*/_{\sim_L}$ is finite
- ► canonical (and minimal) deterministic automaton for L, $\Sigma^*/_{\sim_L}$ as set of states

- $u \sim_L v$ if: for all $w \in \Sigma^*$, $uw \in L$ iff $vw \in L$
- \blacktriangleright u and v have the same "effect" on continuations w
- Myhill-Nerode's Thm: L is regular iff $\Sigma^*/_{\sim L}$ is finite
- \blacktriangleright canonical (and minimal) deterministic automaton for L, $\Sigma^*/_{\sim_I}$ as set of states

Goal

Extend Myhill-Nerode's theorem to classes of transductions

Sequential transductions (Choffrut) Refinement of the MN congruence. Two ideas

1. produce asap: $F(u) = LCP\{f(uw) \mid uw \in dom(f)\}$

Sequential transductions (Choffrut) Refinement of the MN congruence.

Two ideas

- 1. produce asap: $F(u) = LCP\{f(uw) \mid uw \in dom(f)\}$
- 2. $u \sim_f v$ if 2.1 $u \sim_{dom(f)} v$ 2.2 $F(u)^{-1}f(uw) = F(v)^{-1}f(vw)$ $\forall w \in u^{-1}dom(f)$

Sequential transductions (Choffrut) Refinement of the MN congruence.

Two ideas

- 1. produce asap: $F(u) = LCP\{f(uw) \mid uw \in dom(f)\}$
- 2. $u \sim_f v$ if 2.1 $u \sim_{dom(f)} v$ 2.2 $F(u)^{-1}f(uw) = F(v)^{-1}f(vw)$ $\forall w \in u^{-1}dom(f)$

Theorem (Choffrut)

f is sequential iff \sim_f has finite index

 \sim_f is a right congruence \rightsquigarrow canonical and minimal transducer !

Transitions:
$$[u] \xrightarrow{\sigma | F(u)^{-1} F(u\sigma)} [u\sigma]$$

Introduction	Transducers	Logic	Algebra	New logic	Summary

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

• but sequential modulo look-ahead information $\mathcal{I} = \{a, b, \epsilon\}$.

 $a\,b\,b\,a\,a\,a\,a\,b\,b\,b\,a\,b$

Introduction	Transducers	Logic	Algebra	New logic	Summary

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

▶ but sequential modulo *look-ahead information* $\mathcal{I} = \{a, b, \epsilon\}$.

a b b a a a a b b b b a b b

Introduction	Transducers	Logic	Algebra	New logic	Summary

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

▶ but sequential modulo *look-ahead information* $\mathcal{I} = \{a, b, \epsilon\}$.

a b b a a a a b b b b a b b b b a b

Introduction	Transducers	Logic	Algebra	New logic	Summary

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

• but sequential modulo look-ahead information $\mathcal{I} = \{a, b, \epsilon\}$.

a b b a a a a b b b b a b a b

Introduction	Transducers	Logic	Algebra	New logic	Summary

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

▶ but sequential modulo *look-ahead information* $\mathcal{I} = \{a, b, \epsilon\}$.

a b b a a a a b b b b a b

Introduction	Transducers	Logic	Algebra	New logic	Summary

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

▶ but sequential modulo look-ahead information $\mathcal{I} = \{a, b, \epsilon\}$.

Introduction	Transducers	Logic	Algebra	New logic	Summary

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

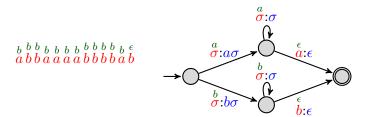
• but sequential modulo look-ahead information $\mathcal{I} = \{a, b, \epsilon\}$.

Introduction	Transducers	Logic	Algebra	New logic	Summary

Rational transductions are almost sequential

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

• but sequential modulo look-ahead information $\mathcal{I} = \{a, b, \epsilon\}$.

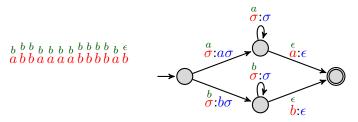


Introduction	Transducers	Logic	Algebra	New logic	Summary

Rational transductions are almost sequential

• $f_{sw}: u\sigma \mapsto \sigma u$ is not sequential

▶ but sequential modulo *look-ahead information* $\mathcal{I} = \{a, b, \epsilon\}$.



- ▶ look-ahead information: $\mathcal{L}: \Sigma^* \to \mathcal{I}$
- ► $f[\mathcal{L}]$: f with input words extended with look-ahead information

Introduction	Transducers	Logic	Algebra	New logic	Summary
Results					
1000 0100					

Theorem (Elgot, Mezei, 65)

f is rational iff $f[\mathcal{L}]$ is sequential, for some finite look-ahead information \mathcal{L} computable by a right sequential transducer. Original statement: $RAT = SEQ \circ RightSEQ$. Introduction Transducers Logic Algebra New logic Summary Results

Theorem (Elgot, Mezei, 65)

f is rational iff $f[\mathcal{L}]$ is sequential, for some finite look-ahead information \mathcal{L} computable by a right sequential transducer. Original statement: $RAT = SEQ \circ RightSEQ$.

Reutenauer, Schützenberger, 91

- canonical look-ahead given by a congruence \equiv_f
- identify suffixes with a 'bounded' effect on the transduction of prefixes
- characterization of rational transductions
 - f is rational
 - ▶ \equiv_f has finite index and $f[\equiv_f]$ is sequential
 - ▶ \equiv_f and $\sim_{f[\equiv_f]}$ have finite index.

New logic

A canonical look-ahead

Identify suffixes with a 'bounded' effect on prefixes.

Equivalent suffixes

▶ $u \equiv_f v$ if $\{d(f(wu), f(wv)) \mid wu, wv \in dom(f)\}$ is finite¹.

A canonical look-ahead

Identify suffixes with a 'bounded' effect on prefixes. Equivalent suffixes

- ▶ $u \equiv_f v$ if $\{d(f(wu), f(wv)) \mid wu, wv \in dom(f)\}$ is finite¹.
- ▶ \equiv_f is a left congruence

A canonical look-ahead

Identify suffixes with a 'bounded' effect on prefixes. Equivalent suffixes

- ▶ $u \equiv_f v$ if $\{d(f(wu), f(wv)) \mid wu, wv \in dom(f)\}$ is finite¹.
- $\blacktriangleright \equiv_f$ is a left congruence
- ▶ look-ahead: return the class of the suffix.

Ex: For f_{sw} , $\Sigma^*/_{\equiv} = \{\{ua \mid u \in \Sigma^*\}, \{ub \mid u \in \Sigma^*\}, \{\epsilon\}\}.$

A canonical look-ahead

Identify suffixes with a 'bounded' effect on prefixes. Equivalent suffixes

- ▶ $u \equiv_f v$ if $\{d(f(wu), f(wv)) \mid wu, wv \in dom(f)\}$ is finite¹.
- $\blacktriangleright \equiv_f$ is a left congruence
- ▶ look-ahead: return the class of the suffix.

Ex: For f_{sw} , $\Sigma^*/_{\equiv} = \{\{ua \mid u \in \Sigma^*\}, \{ub \mid u \in \Sigma^*\}, \{\epsilon\}\}.$ Original Result

- ▶ canonical machine called bimachine
- $\blacktriangleright \text{ bimachine} = \text{left DFA} + \text{right DFA} + \text{output function}$
- right automaton = look-ahead

Definability problems

Rational Transductions

Given f defined by T, is it definable by some \mathcal{C} -transducer ?

- ► sufficient conditions on C to get decidability (F., Gauwin, Lhote, LICS'16)
- ▶ includes aperiodic congruences: decidable FO-definability
- ▶ even PSPACE-C (F., Gauwin, Lhote, FSTTCS'16)

Regular Transductions

 existence of a canonical transducer if *origin* is taken into account (Bojanczyk, ICALP'14)

$$a^{\prime} a^{\prime} a^{\prime} a \mapsto a a a \neq a a^{\prime} a^{\prime} a^{\prime} \mapsto a a a$$

 \blacktriangleright decidable FO-definability with origin, open without

A new logic for transductions joint with Luc Dartois and Nathan Lhote

Algebra Transducers Logic New logic Summar MOTIVATIONS - specify properties of transductions in a high-level formalism : a logic - decidable model-checking a transduer (NFT, 2DFT, ...) a fomula 24/33

Transducers Logic Algebra New logic Summar MOTIVATIONS - specify properties of transductions in a high-level formation : a logic decidable model-checking a transduer a fomula (NFT, 2DFT, ...) $\begin{bmatrix} T \end{bmatrix} \subseteq \begin{bmatrix} Q \end{bmatrix}$ a transduction $\Sigma_{\rightarrow}^* \Sigma^*$ a relation $\subseteq \Sigma_{\times}^* \Sigma^*$ (function) 25/33

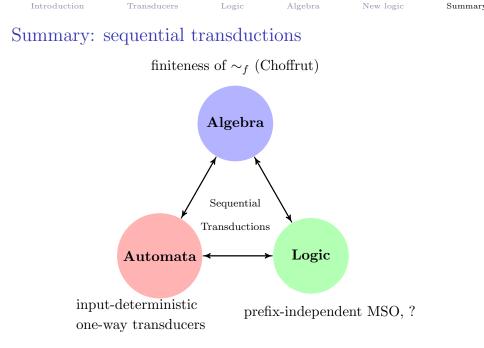
Introduction Transducers Logic Algebra New logic Summar EXAMPLES "there exists at least one 2' in the output" $q'(u, v_1 a v_2) | u, v_1, v_2 \in \Sigma^* \}$ "every request is processed exactly one " { (Γin - Γik , g π(in) -- g π(ik)) / π permutation } more generally : shuffle $\left\{ \left(\sigma_{1} \dots \sigma_{n}, \sigma_{\pi(n)} \dots \sigma_{\pi(n)} \right) \mid \pi \text{ permetation } \right\}$

Transducers Algebra Logic New logic Summar NON-DETERMINISTIC MSOT (NMSOT) COURCELLE use second-order parometers X, ..., Xu { (u, v) / v is a subursed of u, 11/ even } a c a S C Pdan (X) = even (X) $\mathcal{P}_{S}(x,y|X) = x,y \in X \land x < y$ 17 (32EX × < 24y)

Transducers Algebra Logic New logic Summar FROM NMSOT to NEW LOGIC NMSOT is not substactory as a specification language * too "operational" * the previous examples are not NMSOT IDEA : SEE TRANSDUCTIONS AS SINGLE STRUCTURES WITH ORIGIN a ongin a C ourput Predicates: Sin, Sout, or

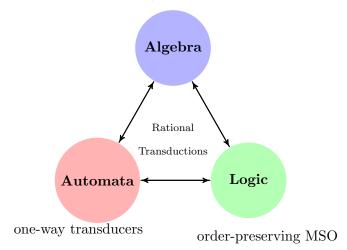
Transducers Algebra Logic New logic Summar EXAMPLES " there exists at least one a " in the output " Jour a(x) shuffle BIJ(0) 1 A V''x Your o(xiy) ~ > o(y) res (x) LABPRES (0) identity BIJ (=) ~ LABPRES(=) ~ ORDER PRES (=) where ORDER PRES(o) = $\forall x, x' \forall y, y' \sigma(x', y') \rightarrow \gamma \leq \sigma r \gamma'$ × Six'

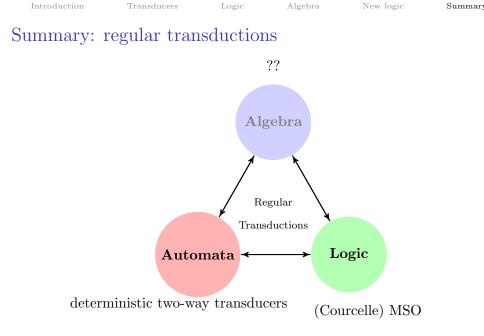
Transducers Algebra New logic Logic Summar RESULTS ON MSO[< in , < wy , &] Devidable model-checking of 2DFT [Bojanczyk' 17] Undecidable satisfiability y [Nathon's talk] Devidable fragment for satisfieb. City)



Summary: rational transductions

finiteness of \equiv_f and $\sim_{f[\equiv_f]}$ (Reutenauer, Schützenberger)



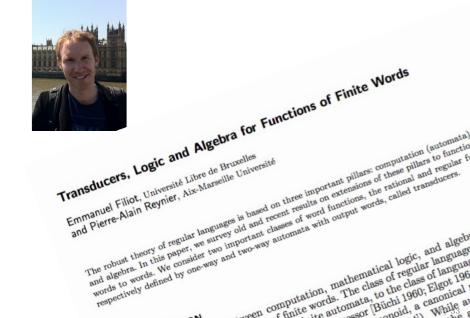




- ▶ AC^0 transductions (Cadilhac,Krebs,Ludwig,Paperman,15)
- variants of two-way transducers (Guillon, Choffrut, 14,15,16), (Carton, 12) (McKenzie, Schwentick, Thérien, Vollmer, 06)
- learning rational transductions (Boiret, Lemay, Niehren, 12)
- inclusion, equivalence and uniformisation problems for rational transductions with "similar origins" (F., Jecker, Löding, Winter, 16)
- ▶ non-determinism
- ▶ infinite words, nested words, trees

Introduction Transducers Logic Algebra New logic Summary

SIGLOG News 9th



Introduction Transducers Logic Algebra New logic Summary

SIGLOG News 9th

