An algebraic characterization of unary two-way transducers

Christian Choffrut¹, Bruno Guillon^{1,2}

¹LIAFA, Université Paris Diderot, Paris 7
²Dipartimento di Informatica, Università degli Studi di Milano

July 2, 2014

Two-way automaton over Σ

$$(Q, \Sigma, I, F, \delta) \longleftarrow$$



Two-way automaton over $\boldsymbol{\Sigma}$



Two-way transducer over Σ , Γ

$$(Q, \Sigma, I, F, \delta) \xleftarrow[]{} (Q, \Sigma, I, F, \delta) \xleftarrow[]{} transition set: \subseteq Q \times \overline{\Sigma} \times \{-1, 0, +1\} \times Q$$



Two-way transducer over Σ , Γ

































$$\Sigma = \Gamma = \{a, b\}$$



accepts:
$$\{(w, w) \mid w \in \Sigma^*\}$$













 $\Sigma = \Gamma = \{a\}$



accepts: $\{(a^n, a^n) \mid n \in \mathbb{N}\}$







back to ⊳













 $R\subseteq \Sigma^*\times \Gamma^*$

Relations are



 $R\subseteq \mathbf{\Sigma}^*\times\mathbf{\Gamma}^*$

the image of $u \in \Sigma^*$ is $R(u) = \{v \mid (u, v) \in R\}$





$$R\subseteq \mathbf{\Sigma}^*\times\mathbf{\Gamma}^*$$

the image of $u \in \Sigma^*$ is $R(u) = \{v \mid (u, v) \in R\}$

$$R: \begin{array}{ccc} \Sigma^* & \to & \mathcal{P}(\Gamma^*) \\ u & \mapsto & R(u) \end{array}$$

RelationsareSeries $R \subseteq \Sigma^* \times \Gamma^*$ $s \in \mathcal{P}(\Gamma^*) \langle \langle \Sigma^* \rangle \rangle$

the image of $u \in \Sigma^*$ is $R(u) = \{v \mid (u, v) \in R\}$ the coefficient of $u \in \Sigma^*$ is $\langle s, u \rangle = R(u)$

$$R: \begin{array}{ccc} \Sigma^* & \to & \mathcal{P}(\Gamma^*) \\ u & \mapsto & R(u) \end{array}$$

Relations	are	Series
$R \subseteq \Sigma^* \times \Gamma^*$		$s\in \mathcal{P}(\Gamma^*)\langle\langle\Sigma^* angle angle$
	Examples	
	$\Sigma = \Gamma = \{a, b\}$	
$R = \left\{ (w, a^{ w _a}) \right\}$		$s = \sum_{w \in \Sigma^*} \left\{ a^{ w _a} \right\} w$
$R = \{(w, ww) \mid w \in \Sigma^*\}$		$s = \sum_{w \in \Sigma^*} \{ww\}w$
	$\Sigma = \Gamma = \{a\}$	
$R = \left\{ (a^n, a^{kn}) \mid k, n \in \mathbb{N} \right\}$		$s = \sum_{a^n \in \Sigma^*} \left\{ a^{kn} \mid k \in \mathbb{N} \right\} a^n$

Rational operations...



 $s+t=\sum_{u\in\Sigma^*}\left(\langle s,u
angle\cup\langle t,u
angle
ight)u$

Rational operations...

Sum: $s + t = \sum_{u \in \Sigma^*} (\langle s, u \rangle \cup \langle t, u \rangle) u$

• Cauchy product: $s \cdot t = \sum_{u \in \Sigma^*} \sum_{u_1 u_2 = u} \langle s, u_1 \rangle \langle t, u_2 \rangle u$

Rational operations...

► Sum:
$$s + t = \sum_{u \in \Sigma^*} (\langle s, u \rangle \cup \langle t, u \rangle) u$$

- ► Cauchy product: $s \cdot t = \sum_{u \in \Sigma^*} \sum_{u_1 u_2 = u} \langle s, u_1 \rangle \langle t, u_2 \rangle u$
- ► Kleene star: $s^* = \sum_{u \in \Sigma^*} \sum_{u_1 u_2 \cdots u_n = u} \langle s, u_1 \rangle \langle s, u_2 \rangle \cdots \langle s, u_n \rangle u$

Rational operations are one-way natural operations






























































Definition

The class of Rational Series is the smallest class of series

Definition

The class of Rational Series is the smallest class of series

that contains Polynomials;

Definition

The class of Rational Series is the smallest class of series

- that contains Polynomials;
- and which is closed under rational operations.

Definition

The class of Rational Series is the smallest class of series

- that contains Polynomials;
- and which is closed under rational operations.

Theorem

One-way transducers accepts exactly the class of rational series.

Hadamard operations...

Hadamard product:

$$s \oplus t = \sum_{u \in \Sigma^*} \langle s, u \rangle \cdot \langle t, u \rangle u$$

Hadamard operations...

Hadamard product:

$$s \oplus t = \sum_{u \in \Sigma^*} \langle s, u \rangle \cdot \langle t, u \rangle u$$

Hadamard star:

$$s^{\mathsf{H}\star} = \sum_{n \in \mathbb{N}} \underbrace{s \oplus s \oplus \dots \oplus s}_{n \text{ times}} = \sum_{u \in \Sigma^*} \langle s, u \rangle^* u$$
Hadamard product: **s** (**b t**)



Hadamard product: **s** (**b t**)



Hadamard product: **s** (**b t**)



Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$



Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$



Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$



Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$



Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



11/26

Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$




Hadamard product: **s** (**b t**)



Hadamard product: **s** (**b t**)



Hadamard product: **s** (**b t**)



Hadamard product: **s** (**b t**)





Hadamard product: **s** (**b t**)











Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: **s** (**b t**)



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



Hadamard product: $s \oplus t$



$$\Sigma = \Gamma = \{a, b\}$$











$$\Sigma = \Gamma = \{a\}$$



$$\left(\Sigma = \Gamma = \{a\}\right)$$





back to \triangleright

$$\Sigma = \Gamma = \{a\}$$






Hadamard-rational

Definition $s \in \mathcal{P}(\Gamma^*)\langle\langle \Sigma^* \rangle\rangle$ is H-Rat if $s = \sum_i \alpha_i \oplus \beta_i^{\mathsf{H}\star}$

where the sum is finite and α_i s and β_i s are rational.

Hadamard-rational

Definition $s\in \mathcal{P}(\Gamma^*)\langle\langle\Sigma^*
angle
angle$ is H-Rat if $s=\sum\limits_i lpha_i\opluseta_i^{\mathsf{H}\star}$

where the sum is finite and α_i s and β_i s are rational.

Property



Hadamard-rational

Definition $s\in \mathcal{P}(\Gamma^*)\langle\langle\Sigma^*
angle
angle$ is H-Rat if $s=\sum\limits_i lpha_i\opluseta_i^{\mathsf{H}\star}$

where the sum is finite and α_i s and β_i s are rational.



Main Result

Definition $s \in \mathcal{P}(\Gamma^*)\langle\langle \Sigma^* \rangle\rangle$ is H-Rat if $S = \sum_i \alpha_i \oplus \beta_i^{\mathsf{H}\star}$ where the sum is finite and α_i s and β_i s are rational.

Main Result

Definition $s \in \mathcal{P}(\Gamma^*)\langle\langle \Sigma^* \rangle\rangle$ is H-Rat if $S = \sum_i \alpha_i \oplus \beta_i^{H\star}$ where the sum is finite and α_i s and β_i s are rational.

Theorem Unary two-way transducers accepts exactly H-Rat series. Analogy with Probabilistic Automata

Theorem (Anselmo, Bertoni, 1994)

Acceptation probability of two-way finite automata is of the form:

$$\tau(w) = \alpha(w) \times \frac{1}{\beta(w)}$$

where α and β are rational series of $\mathbb{Q}\langle\langle \Sigma^* \rangle\rangle$.

Theorem (Engelfriet, Hoogeboom)

Two-way transducers versus MSO logic

Two-way transducers	versus	MSO logic	
functional		functions	

Two-way transducers	versus	MSO logic	
functional		functions	

Two-way transducers	versus	MSO logic	
functional		functions	
general		relations	

Two-way transducers	versus	MSO logic	
functional		functions	
general	incomparable	relations	

Theorem (Engelfriet, Hoogeboom)

Two-way transducers	versus	MSO logic	
functional		functions	
general	incomparable	relations	

Theorem (Filiot, Gauwin, Reinier, Servais)

Theorem (Engelfriet, Hoogeboom)

Two-way transducers	versus	MSO logic	
functional		functions	
general	incomparable	relations	

Theorem (Filiot, Gauwin, Reinier, Servais)



Theorem (Engelfriet, Hoogeboom)



Theorem (Filiot, Gauwin, Reinier, Servais)



Theorem (Engelfriet, Hoogeboom)

Two-way transducers	versus	MSO logic
functional		functions
general	incomparable	relations

Theorem (Filiot, Gauwin, Reinier, Servais)



Theorem (Engelfriet, Hoogeboom)

Two-way transducers	versus	MSO logic
functional		functions
general	incomparable	relations

Theorem (Filiot, Gauwin, Reinier, Servais) unary alphabets?



























Get around the problems...

Consider only loop-free runs;

Get around the problems...

Consider only loop-free runs;

• Consider the case $\Gamma = \{a\}$, or parikh-equivalence.

Particular transducers

F^{*} is commutative

Theorem

For any deterministic or functional transducer

there exists an equivalent one-way transducer.

Particular transducers * is commutative Theorem For any deterministic or functional transducer there exists an equivalent one-way transducer. Theorem For any transducer accepting a relation R, \exists a one-way transducer accepting a rational uniformization of R Particular transducers * is commutative Theorem For any deterministic or functional transducer there exists an equivalent one-way transducer. Theorem For any transducer accepting a relation R, \exists a one-way transducer accepting a rational uniformization of R

Input unary case



Lemma

Central loops of a two-way transducer produce finitely many

rational output languages.



Lemma

Central loops of a two-way transducer produce finitely many

rational output languages.

We can take into account central loops in one-way simulation.

Get around the problems...



Consider only loop-free runs,

• Consider the case $\Gamma = \{a\}$, or parikh-equivalence.
Get around the problems...



Consider particular parts of run: hits (from border to border)

• Consider the case $\Gamma = \{a\}$, or parikh-equivalence.



Hit: a border to border run

- Hit: a border to border run
 - ▶ reading *u*



- Hit: a border to border run
 - reading u
 - outputing v



- Hit: a border to border run
 - reading u
 - outputing v
 - no visit to endmarkers



- Hit: a border to border run
 - reading u
 - outputing v
 - no visit to endmarkers

define a relation R_{b_i} , b_j

- Hit: a border to border run
 - reading u
 - outputing v
 - no visit to endmarkers







Given:

- a b_0 to b_x hit over u producing v_0 ;
- and a b_x to b_1 hit over u producing v_1

we may compose them into a b_0 to b_1 run over u producing v_0v_1 .

Given:

- a b_0 to b_x hit over u producing v_0 ;
- and a b_{x} to b_{1} hit over u producing v_{1}

we may compose them into a b_0 to b_1 run over u producing v_0v_1 .

double-hit relations are:

$$\mathsf{R}^{(2)}_{b_0,b_1} = \bigcup_{\substack{b_x \in Q \times \{ \triangleright, \triangleleft \}}} \mathsf{R}_{b_0,b_x} \oplus \mathsf{R}_{b_x,b_1}$$

Given:

- a b_0 to b_x hit over u producing v_0 ;
- and a b_x to b_1 hit over u producing v_1

we may compose them into a b_0 to b_1 run over u producing v_0v_1 .

double-hit relations are:

$$\mathsf{R}^{(2)}_{b_0,b_1} = \bigcup_{\substack{b_{\mathsf{x}} \in Q \times \{ \triangleright, \triangleleft \}}} \mathsf{R}_{b_0,b_{\mathsf{x}}} \oplus \mathsf{R}_{b_{\mathsf{x}},b_1}$$

coefficient (b_0, b_1) of $HIT \oplus HIT$.

Given:

- a b_0 to b_x hit over u producing v_0 ;
- and a b_x to b_1 hit over u producing v_1

we may compose them into a b_0 to b_1 run over u producing $v_0 v_1$.

triple-hit relations are:

$$R^{(3)}_{b_0,b_1} = \bigcup_{b_{x_1},b_{x_2} \in Q \times \{ \triangleright, \triangleleft \}} R_{b_0,b_{x_1}} \oplus R_{b_{x_1},b_{x_2}} \oplus R_{b_{x_2},b_1}$$

coefficient (b_0, b_1) of $HIT \oplus HIT \oplus HIT$.

Given:

- a b_0 to b_x hit over u producing v_0 ;
- and a b_x to b_1 hit over u producing v_1

we may compose them into a b_0 to b_1 run over u producing $v_0 v_1$.

multi-hit relations are:

$$R_{b_0,b_1}^{(\mathsf{H}\star)} = \bigcup_{n \in \mathbb{N}} \bigcup_{b_{\mathsf{x}_1}, \dots, b_{\mathsf{x}_n}} R_{b_0, b_{\mathsf{x}_1}} \oplus \dots, R_{b_{\mathsf{x}_n}, b_1}$$

coefficient (b_0, b_1) of HIT^{H*} .

An accepting run is a particular composition of hits.

An accepting run is a particular composition of hits.

Look at coefficients $R_{b_0,b_1}^{(H\star)}$ of $HIT^{H\star}$ such that:

- b₀ corresponds to the initial configuration
- b_1 corresponds to some accepting configuration

An accepting run is a particular composition of hits.

Look at coefficients $R_{b_0,b_1}^{(H\star)}$ of $HIT^{H\star}$ such that:

- b₀ corresponds to the initial configuration
- b_1 corresponds to some accepting configuration

 $R = \bigcup_{b_1 \text{ accepting}} R_{b_0, b_1}^{(H\star)}$ is the relation accepted by the transducer.

An accepting run is a particular composition of hits.

Look at coefficients $R_{b_0,b_1}^{(H\star)}$ of $HIT^{H\star}$ such that:

- b₀ corresponds to the initial configuration
- b_1 corresponds to some accepting configuration

 $R = \bigcup_{b_1 \text{ accepting}} R_{b_0, b_1}^{(H\star)}$ is the relation accepted by the transducer.

Theorem R is in H-Rat

(by closure properties of H-Rat)

Theorem

Unary two-way transducers accept series: $s = \sum_{i} \alpha_{i} \oplus \beta_{i}^{H\star}$.

Theorem

Unary two-way transducers accept series: $s = \sum_{i} \alpha_{i} \oplus \beta_{i}^{H\star}$.

Corollary Every unary two-way transducer can be made sweeping.

Theorem

Unary two-way transducers accept series: $s = \sum_{i} \alpha_{i} \oplus \beta_{i}^{H\star}$.

Corollary Every unary two-way transducer can be made sweeping.

Theorem



Theorem

Unary two-way transducers accept series: $s = \sum_{i} \alpha_{i} \oplus \beta_{i}^{H\star}$.

Corollary Every unary two-way transducer can be made sweeping.

Theorem



Theorem

Unary two-way transducers accept series: $s = \sum_{i} \alpha_{i} \oplus \beta_{i}^{H\star}$.

Corollary Every unary two-way transducer can be made sweeping.

Theorem



Grazie infinite.