Unambiguity in Transducer Theory

Emmanuel Filiot   Sarah Winter

Université libre de Bruxelles (ULB), Belgium

November 2021
Dagstuhl Seminar on Unambiguity

WHAT IS A (ONE-WAY) TRANSDUCER?

TRANSDUCERS are automata with output.

$\mathcal{A} = (\Sigma, Q, q_0, F, \Delta) \leftarrow$ NFA

Monoid

$\mathcal{T} = (\mathcal{A}, o)$ where $o : \Delta \to (\mathcal{M}, \otimes)$

WHAT IS A (ONE-WAY) TRANSDUCER ?

TRANSDUCERS are automata with output.

$\mathcal{A} = (\Sigma, Q, q_0, F, \Delta) \leftarrow$ NFA

Monoid

$\mathcal{T} = (\mathcal{A}, o)$ where $o : \Delta \to (M, \otimes)$

RELATION $[\![\mathcal{T}]\!] \subseteq \Sigma^* \times M$

WHAT IS A (ONE-WAY) TRANSDUCER ?

TRANSDUCERS are automata with output.

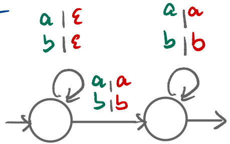
$\mathcal{A} = (\Sigma, Q, q_0, F, \Delta)$ ← NFA

Monoid

$\mathcal{T} = (\mathcal{A}, o)$ where $o : \Delta \to (\mathcal{M}, \otimes)$

RELATION $[\![\mathcal{T}]\!] \subseteq \Sigma^* \times \mathcal{M}$

NFT



NON-DET. FINITE TRANSDUCER

WHAT IS A (ONE-WAY) TRANSDUCER ?

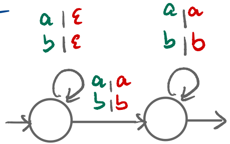
TRANSDUCERS are automata with output.

$\mathcal{A} = (\Sigma, Q, q_0, F, \Delta)$ ← NFA

$\mathcal{T} = (\mathcal{A}, o)$ where $o : \Delta \to (\mathcal{M}, \otimes)$ ← Monoid

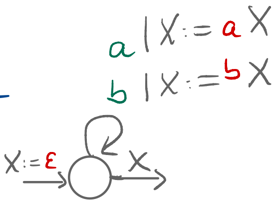
RELATION $[\![\mathcal{T}]\!] \subseteq \Sigma^* \times \mathcal{M}$

NFT



NON-DET. FINITE TRANSDUCER

SST



STREAMING STRING TRANSDUCER

Alur, Černý '10

FUNCTIONAL TRANSDUCERS

- $\mathcal{T} = (\mathcal{A}, o)$ is FUNCTIONAL if $[\![\mathcal{T}]\!]$ is a function.

FUNCTIONAL TRANSDUCERS

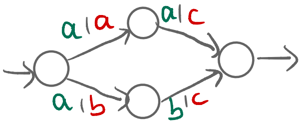
- $\mathcal{T} = (\mathcal{A}, \sigma)$ is FUNCTIONAL if $[\![\mathcal{T}]\!]$ is a function.
- $\mathcal{T} = (\mathcal{A}, \sigma)$ is UNAMBIGUOUS if $\mathcal{A}$ is unambiguous.

FUNCTIONAL TRANSDUCERS

- $\mathcal{T} = (\mathcal{A}, \sigma)$ is FUNCTIONAL if $[\![\mathcal{T}]\!]$ is a function.
- $\mathcal{T} = (\mathcal{A}, \sigma)$ is UNAMBIGUOUS if $\mathcal{A}$ is unambiguous.

EXAMPLES

- unambiguous

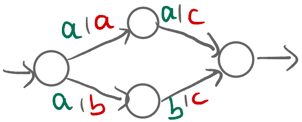


FUNCTIONAL TRANSDUCERS

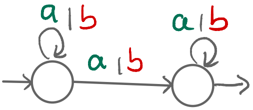
- $\mathcal{T} = (\mathcal{A}, \sigma)$ is FUNCTIONAL if $[\![\mathcal{T}]\!]$ is a function.
- $\mathcal{T} = (\mathcal{A}, \sigma)$ is UNAMBIGUOUS if $\mathcal{A}$ is unambiguous.

EXAMPLES

- unambiguous



- non-deterministic

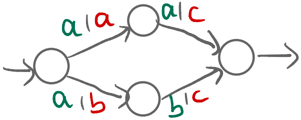


FUNCTIONAL TRANSDUCERS

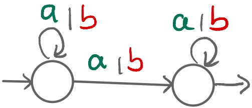
- $\mathcal{T} = (\mathcal{A}, \sigma)$ is FUNCTIONAL if $[\![\mathcal{T}]\!]$ is a function.
- $\mathcal{T} = (\mathcal{A}, \sigma)$ is UNAMBIGUOUS if $\mathcal{A}$ is unambiguous.

EXAMPLES

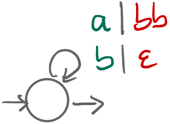
- unambiguous



- non-deterministic



- sequential (= input-deterministic)



FUNCTIONALITY ≡ UNAMBIGUITY

Thm Every function recognized by a transducer $T = (A, \sigma : \Delta \to M)$ can be recognized by an unambiguous one.

FUNCTIONALITY $\equiv$ UNAMBIGUITY

Thm Every function recognized by a transducer $T = (A, \sigma : \Delta \to M)$ can be recognized by an unambiguous one.

EXAMPLES

- RATIONAL FUNCTIONS $\quad M = (\Sigma^*, \cdot)$

  fNFTs $\equiv$ UFTs

FUNCTIONALITY ≡ UNAMBIGUITY

<u>Thm</u> Every function recognized by a transducer $T = (A, \sigma : \Delta \to M)$ can be recognized by an unambiguous one.

EXAMPLES

- RATIONAL FUNCTIONS $\qquad M = (\Sigma^*, \cdot)$

  fNFTs ≡ UFTs

- REGULAR FUNCTIONS $\qquad M = (\text{Substitutions}_X, \circ)$

  fSSTs ≡ USSTs

FUNCTIONALITY ≡ UNAMBIGUITY

Thm Every function recognized by a transducer $T = (A, o : \Delta \to M)$ can be recognized by an unambiguous one.

EXAMPLES

- RATIONAL FUNCTIONS $\qquad M = (\Sigma^*, \cdot)$

  f NFTs ≡ UFTs

- REGULAR FUNCTIONS $\qquad M = (\text{Substitutions}_X, \circ)$

  f SSTs ≡ USSTs

- functional SUM-automata ≡ unambiguous SUM-automata

  $M = (\mathbb{Z}, +)$

MORE GENERAL RESULT (UNIFORMISATION PROPERTY)

Thm Every relation recognized by a transducer contains a function with equal domain that can be recognized by an unambiguous transducer.

$\forall R\ \exists f:\ f \subseteq R\ \wedge\ dom(f) = dom(R)$

MORE GENERAL RESULT (UNIFORMISATION PROPERTY)

Thm Every relation recognized by a transducer contains a function with equal domain that can be recognized by an unambiguous transducer.

$$\forall R\ \exists f:\ f \subseteq R \wedge \mathrm{dom}(f) = \mathrm{dom}(R)$$

Proof Order the runs. For each input word select the smallest accepting run.

MORE GENERAL RESULT (UNIFORMISATION PROPERTY)

Thm Every relation recognized by a transducer contains a function with equal domain that can be recognized by an unambiguous transducer.

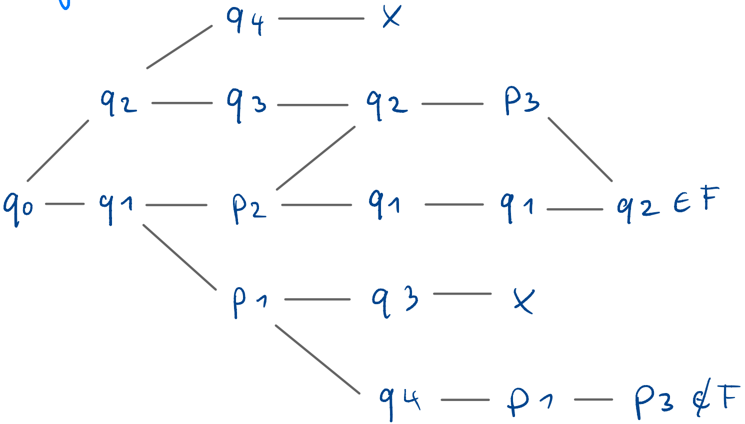
$\forall R \, \exists f : f \subseteq R \wedge dom(f) = dom(R)$

Proof Order the runs. For each input word select the smallest accepting run.

One of many proofs.
(First one: Eilenberg '74)

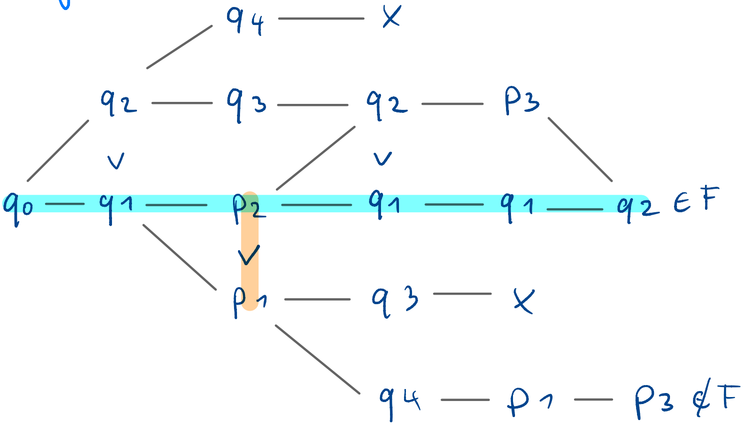
PROOF

Selecting the smallest run via any fixed ordering on the states.



PROOF

Selecting the smallest run via any fixed ordering on the states.



q_0	q_1	p_2	q_1	q_1	$q_2 \in F$
$\emptyset$	$\emptyset$	$\{p_1\}$	$\{q_3, q_4\}$	$\{p_1\}$	$\{p_3\} \cap F = \emptyset$

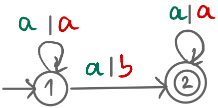
TRANSDUCERS ON INFINITE WORDS

$T = (A, \sigma)$, A with Büchi acceptance.

Thm Every relation recognized by a Büchi-transducer contains a function with equal domain that can be recognized by an unambiguous Büchi-transducer.

Proof Order the runs. Select the smallest accepting run.

⚠ Not every state ordering works ⚠

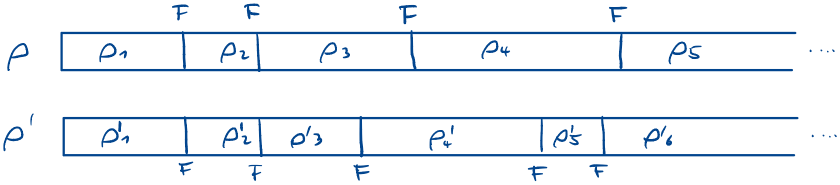


$2 < 1$: For every accepting run there exists a smaller accepting run.

PROOF IDEA: HOW TO ORDER RUNS

Choffrut & Grigorieff '99

Cut runs into segments at final states



- Segment $\rho_i < \rho'_i$ if $\rho_i <_{llex} \rho'_i$ for some fixed ordering on Q

  length-lexicographic

- Run $\rho < \rho'$ if there is some i s.t. $\rho_i < \rho'_i$ and for all $j < i$: $\rho_j = \rho'_j$.

SUCCINTNESS OF UNAMBIGUOUS TRANSDUCERS

Thm NFA $<_{\exp}$ UFA Schmidt'77

SUCCINTNESS OF UNAMBIGUOUS TRANSDUCERS

Thm NFA $<_{exp}$ UFA Schmidt '77

- $L_n = \{ u \# v \mid |u| = |v| = n, u \neq v, u, v \in \{a,b\}^* \}$

SUCCINTNESS OF UNAMBIGUOUS TRANSDUCERS

Thm   NFA $<_{exp}$ UFA Schmidt '77

- $L_n = \{ u\#v \mid |u| = |v| = n,\ u \neq v,\ u, v \in \{a,b\}^* \}$
- Smallest NFA is of size $O(n)$: Guess i s.t. $u_i \neq v_i$
- Up to n different accepting runs

SUCCINTNESS OF UNAMBIGUOUS TRANSDUCERS

Thm NFA $<_{\exp}$ UFA Schmidt '77

- $L_n = \{ u\#v \mid |u| = |v| = n,\ u \neq v,\ u, v \in \{a,b\}^* \}$
- Smallest NFA is of size $O(n)$: Guess i s.t. $u_i \neq v_i$
- Up to n different accepting runs

$\Rightarrow$ f NFT $<_{\exp}$ UFT

SUCCINTNESS OF UNAMBIGUOUS TRANSDUCERS

<u>Thm</u> NFA $<_{exp}$ UFA Schmidt '77

- $L_n = \{u\#v \mid |u| = |v| = n,\ u \neq v,\ u, v \in \{a,b\}^*\}$
- Smallest NFA is of size $O(n)$: Guess i s.t. $u_i \neq v_i$
- Up to n different accepting runs

$\Rightarrow$ fNFT $<_{exp}$ UFT, but

OPEN

$f\text{NFT}_{\substack{\text{total}\\\text{domain}}} <_{exp} \text{UFT}_{\substack{\text{total}\\\text{domain}}}$?

IS UNAMBIGUITY NECESSARY?

EXAMPLES (from $M = \Sigma^*$)

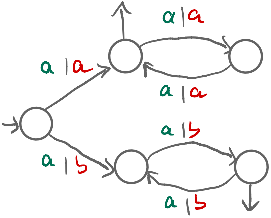
$$f_1 : \underset{n \geq 1}{a^n} \mapsto \begin{cases} a^n & n \text{ odd} \\ b^n & n \text{ even} \end{cases}$$

IS UNAMBIGUITY NECESSARY?

EXAMPLES (from $M = \Sigma^*$)

$$f_1 : a^n \mapsto \begin{cases} a^n & n \text{ odd} \\ b^n & n \text{ even} \end{cases}$$

$n \geq 1$

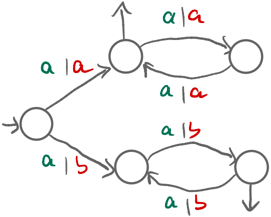


IS UNAMBIGUITY NECESSARY?

EXAMPLES (from $M = \Sigma^*$)

$$f_1 : a^n \mapsto \begin{cases} a^n & n \text{ odd} \\ b^n & n \text{ even} \end{cases}$$

$n \geq 1$



$$f_2 : uba^n \mapsto ub$$

$u \in \{a,b\}^*$

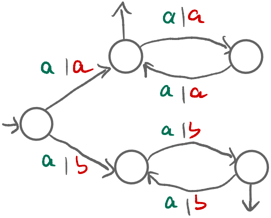
$n \geq 0$

IS UNAMBIGUITY NECESSARY?

EXAMPLES (from $M = \Sigma^*$)

$$f_1 : a^n \mapsto \begin{cases} a^n & n \text{ odd} \\ b^n & n \text{ even} \end{cases}$$

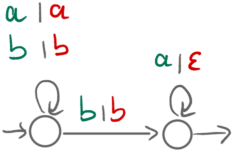
$n \geq 1$



$$f_2 : uba^n \mapsto ub$$

$u \in \{a, b\}^*$

$n \geq 0$



IS UNAMBIGUITY NECESSARY?

f_3: remove unnecessary white spaces

$$a\sqcup\sqcup a\sqcup aa\sqcup \mapsto a\sqcup a\sqcup aa$$

IS UNAMBIGUITY NECESSARY?

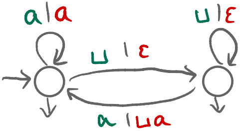
f_3: remove unnecessary white spaces

a⊔⊔a⊔aa⊔ ⟼ a⊔a⊔aa

IS UNAMBIGUITY NECESSARY?

f_3: remove unnecessary white spaces

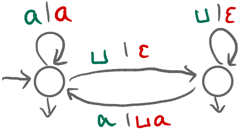
$a\sqcup\sqcup a\sqcup aa\sqcup \mapsto a\sqcup a\sqcup aa$



IS UNAMBIGUITY NECESSARY?

f_3: remove unnecessary white spaces

$a ⊔⊔⊔ a ⊔ ⊔ aa ⊔ \mapsto a⊔a⊔aa$



Thm Given a fNFT $\mathcal{T}$, it is decidable in PTIME whether $\mathcal{T}$ is equivalent to some sequential (= input-deterministic) transducer.

Béal, Carton, Prieur, Sakarovitch '03

REMOVING UNAMBIGUITY WITH TWO-WAYNESS

$f_2 : uba^n \mapsto ub$

$u \in \{a, b\}^*$

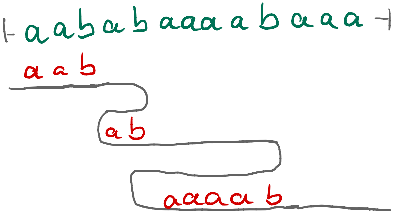
$n \in \mathbb{N}$

REMOVING UNAMBIGUITY WITH TWO-WAYNESS

$f_2 : uba^n \mapsto ub$

$u \in \{a,b\}^*$

$n \in \mathbb{N}$

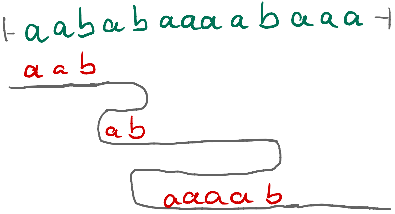


REMOVING UNAMBIGUITY WITH TWO-WAYNESS

$f_2 : uba^n \mapsto ub$

$u \in \{a,b\}^*$

$n \in \mathbb{N}$



Thm For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

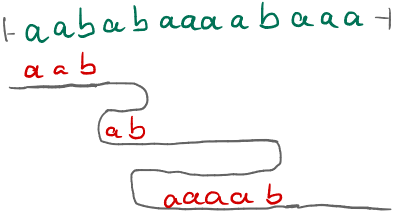
Chytil, Jákl '77

REMOVING UNAMBIGUITY WITH TWO-WAYNESS

$f_2 : u b a^n \mapsto ub$

$u \in \{a, b\}^*$

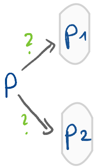
$n \in \mathbb{N}$



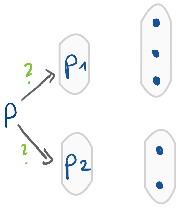
Thm For every unambiguous and co-deterministic one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

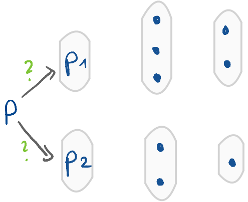
HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67



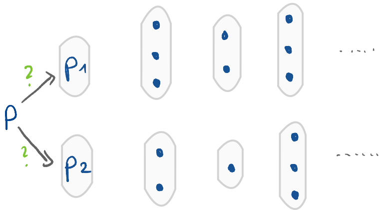
HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67



HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67

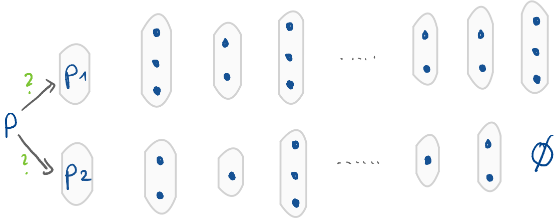


HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67

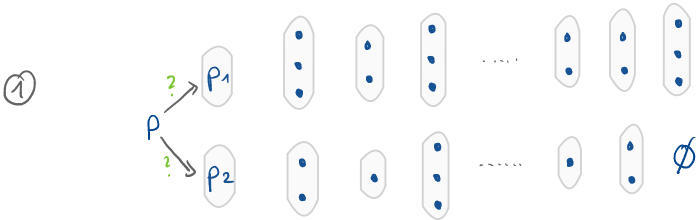


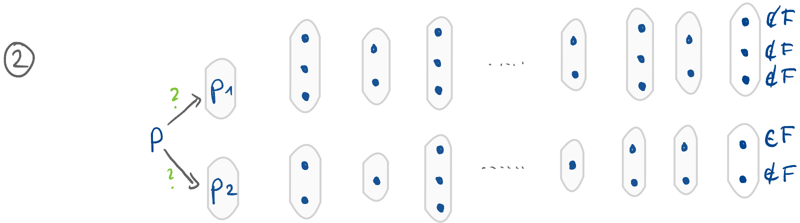
HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67

①

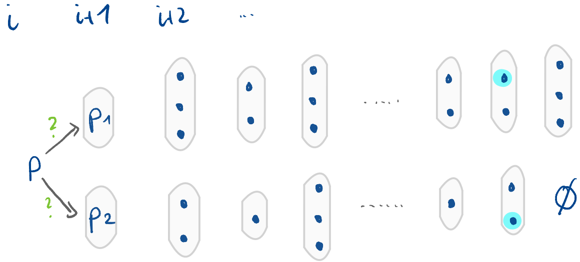


HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67



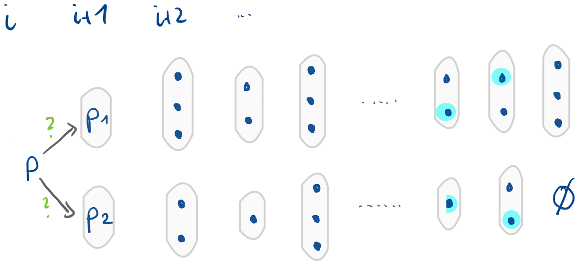


HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67



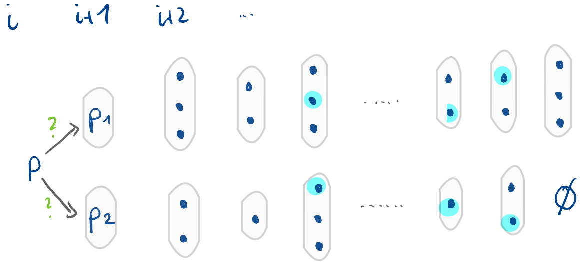
- Bags are disjoint because $\mathcal{A}$ is a co-det. FA
- ensures that the runs meet at position i

HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67



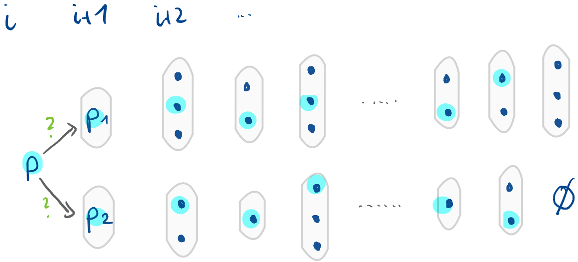
- Bags are disjoint because $\mathcal{A}$ is a co-det. FA
- ensures that the runs meet at position i

HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67



- Bags are disjoint because $\mathcal{A}$ is a co-det. FA
- ensures that the runs meet at position i

HEART OF THE CONSTRUCTION: HOPCROFT-ULLMAN '67



- Bags are disjoint because $\mathcal{A}$ is a co-det. FA
- ensures that the runs meet at position i

GENERAL CASE

<u>Thm</u> For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

GENERAL CASE

<u>Thm</u> For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

<u>Proof</u>

- Every rational function f can be expressed as $f = f_2 \circ f_1$, f_2 is left-sequential, f_1 is right-sequential. Elgot, Mezei '65

GENERAL CASE

Thm For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

Proof

- Every rational function f can be expressed as $f = f_2 \circ f_1$, f_2 is left-sequential, f_1 is right-sequential. Elgot, Mezei '65

- $$f = \overset{\text{DET}}{f_2} \circ \overset{\text{coDET}}{f_1}$$

GENERAL CASE

Thm For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

Proof

- Every rational function f can be expressed as $f = f_2 \circ f_1$, f_2 is left-sequential, f_1 is right-sequential. Elgot, Mezei '65
- $$f = \overset{\text{DET}}{f_2} \circ \overset{\text{coDET}}{f_1} = \overset{\text{DET}}{f_2} \circ \overset{\text{2-DET}}{f_1}$$

GENERAL CASE

Thm For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

Proof

- Every rational function f can be expressed as $f = f_2 \circ f_1$, f_2 is left-sequential, f_1 is right-sequential. Elgot, Mezei '65
- $$f = \underbrace{f_2}_{\text{DET}} \circ \underbrace{f_1}_{\text{coDET}} = \underbrace{f_2}_{\text{DET}} \circ \underbrace{f_1}_{\text{2-DET}}$$
- $\text{RAT} \underset{\text{Elgot Mezei}}{\equiv} \text{DET} \circ \text{coDET} \underset{\text{Hopcroft Ullman}}{\equiv} \text{DET} \circ \text{2DET}$

GENERAL CASE

Thm For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

Proof

- Every rational function f can be expressed as $f = f_2 \circ f_1$, f_2 is left-sequential, f_1 is right-sequential. Elgot, Mezei '65

- $$f = \underbrace{f_2}_{\text{DET}} \circ \underbrace{f_1}_{\text{coDET}} = \underbrace{f_2}_{\text{DET}} \circ \underbrace{f_1}_{\text{2-DET}}$$

- RAT $\equiv$ DET $\circ$ coDET $\equiv$ DET $\circ$ 2DET $\equiv$ 2DET

Elgot Mezei

Hopcroft Ullman

GENERAL CASE

Thm For every unambiguous one-way transducer there is an equivalent deterministic two-way transducer.

Chytil, Jákl '77

Proof

- Every rational function f can be expressed as $f = f_2 \circ f_1$, f_2 is left-sequential, f_1 is right-sequential. Elgot, Mezei '65

- $$f = \overset{\text{DET}}{f_2} \circ \overset{\text{coDET}}{f_1} = \overset{\text{DET}}{f_2} \circ \overset{\text{2-DET}}{f_1}$$

- $$\underset{\text{Elgot Mezei}}{\text{RAT} \equiv} \text{DET} \circ \text{coDET} \underset{\text{Hopcroft Ullman}}{\equiv} \text{DET} \circ \text{2DET} \underset{\wedge}{\equiv} \text{2DET}$$

2WAY $q \xrightarrow{\sigma \mid u, d} q'$

DET $p \xrightarrow{u \mid v} p'$

COMP $(q,p) \xrightarrow{\sigma \mid v, d} (q', p')$

SUMMARY

Thm $f\text{NFT} \equiv \text{UFT} \supsetneq$ sequential FTs

$\cap\!\mid$

2DFT

SUMMARY

<u>Thm</u> $f\,\mathrm{NFT} \equiv \mathrm{UFT} \supsetneq$ sequential FTs

$$\begin{array}{c} f\,\mathrm{NFT} \\ \cap\!\mathrm{I} \\ \mathrm{2DFT} \end{array}$$

<u>Thm</u> $f\,\mathrm{SST} \equiv \mathrm{USST}$

SUMMARY & MORE

Thm

$$\begin{array}{l} f\text{NFT} \equiv \text{UFT} \supsetneq \text{sequential FTs} \\ \quad \cap\!\!| \\ \text{2DFT} \end{array}$$

Thm

$$f\,\underbrace{\text{SST} \equiv \text{USST} \equiv \text{DSST} \equiv \text{2DFT}}_{\text{Alur, Černý '10}}$$

SUMMARY & MORE

Thm $\quad f\text{NFT} \equiv \text{UFT} \supsetneq \text{sequential FTs}$

$$\begin{array}{c} f\text{NFT} \\ \cap\!\!\shortmid \\ \text{2DFT} \end{array}$$

Dartois, Fournier, Jecker, Lhote '17

Thm $\quad \underbrace{f\text{SST} \equiv \text{USST} \equiv \text{DSST} \equiv \text{2DFT}}_{\text{Alur, Černý '10}} \equiv \text{2RFT}$

$\boxed{\text{REVERSIBLE} = \text{det. \& co-det.}}$

SUMMARY & MORE

Thm fNFT $\equiv$ UFT $\supsetneq$ sequential FTs
$\cap$
2DFT

Dartois, Fournier, Jecker, Lhote '17

Thm $\underbrace{f\text{SST} \equiv \text{USST} \equiv \text{DSST} \equiv \text{2DFT}}_{\text{Alur, Černý '10}} \equiv$ 2RFT

REVERSIBLE
= det. & co-det.

Thm Every relation recognizable by a 2NFT contains a function with equal domain that can be recognized by a
| 2DFT De Souza '13
| 2RFT Dartois, Fournier, Jecker, Lhote '17

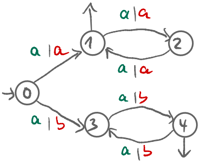
FUNCTION CONCATENATION

Thm Every rational function f can be expressed as $f = f_2 \circ f_1$, f_2 is left-sequential, f_1 is right-sequential. Elgot, Mezei '65

right-seq. (co-det. 1FT): on input $u = u_1 u_2 \dots u_n$ the ith output letter $= \{ (p,q) \mid A : p \xrightarrow{u_i \dots u_n} q \in F \}$

a	a	a	a	a	a	
(1,1)	(0,1)	(1,1)	(0,1)	(1,1)	(0,1)	(1,1)
(0,4)	(2,1)	(0,4)	(2,1)	(0,4)	(2,1)	(4,4)
(4,4)	(3,4)	(4,4)	(3,4)	(0,4)	(3,4)	

$\mathcal{T} = (A, o)$ unambiguous



left-sequential

(1,1)	(0,1)	(1,1)	(0,1)	(1,1)	(0,1)	(1,1)
(0,4)	(2,1)	(0,4)	(2,1)	(0,4)	(2,1)	(4,4)
(4,4)	(3,4)	(4,4)	(3,4)	(0,4)	(3,4)	
ε	b	b	b	b	b	b