A Regular and Complete Notion of Delay for Streaming String Transducers

Sarah Winter

Université libre de Bruxelles (ULB)

joint work with

E. Filiot
ULB

I. Jecker
Warsaw

C. Löding
RWTH

String-to-String Transductions

FUNCTIONAL TRANSDUCTIONS $f : \Sigma^* \to \Sigma^*$

* Delete all a — $aababababb \mapsto bbbb$

* Count the length — $abaab \mapsto c^5$

* Mirror — $abaab \mapsto baaba$

* Copy — $abaab \mapsto abaababaab$

* Remove superfluous whitespaces

  ␣This␣␣is␣some␣␣␣text.␣ $\mapsto$ This␣is␣some␣text.

* Reactive synthesis: grant every request

  $\ulcorner\ g\ \urcorner\ulcorner\ \urcorner g\ \urcorner\ulcorner\ \urcorner g\ \ulcorner\ g\ \urcorner\ulcorner\ \urcorner g$

  $\ulcorner\ \urcorner\ulcorner\ \urcorner\ulcorner\ \ulcorner\ \urcorner\ulcorner \mapsto g\ \urcorner g\ \urcorner g\ g\ \urcorner g$

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

* Infix relation

$\{(popcorn, corn), (popcorn, opc), (popcorn, \varepsilon), \ldots\}$

* Shuffle

$\{(popcorn, pcoprno), (popcorn, copprno), \ldots\}$

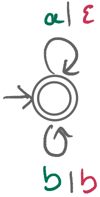
COMPUTATION MODELS
FOR TRANSDUCTIONS

TRANSDUCERS

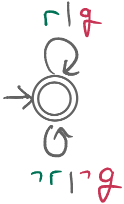
RATIONAL TRANSDUCTIONS

MODEL: finite automata with outputs

* delete all a



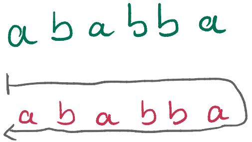
* Reactive synthesis: grant every request



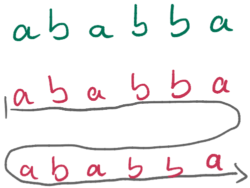
REGULAR TRANSDUCTIONS

MODEL 2-way finite automata with outputs

* mirror



* copy

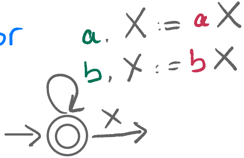


A model without 2-way?

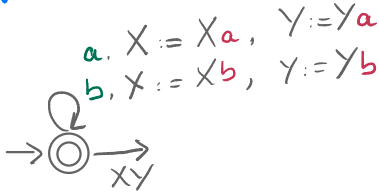
STREAMING STRING TRANSDUCERS (SST)

MODEL (deterministic) finite automata with registers

* mirror



* copy



Here:
copyless SST

~~X := Xa~~
~~Y := XbbY~~

REGULAR TRANSDUCTIONS

MANY MODELS

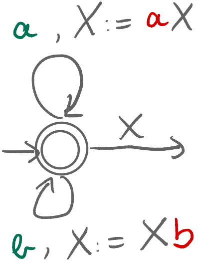
* 2-way derministic transducers (2DFT)
* Courcelle's MSO transducers (MSOT)
* Streaming string transducers (SST) [Alur, Černý '10]
* regular combinators [Alur, Freilich, Raghothaman '14]
* 2-way reversible transducers [Dartois, Fournier, Jecker, Lhote, '17]

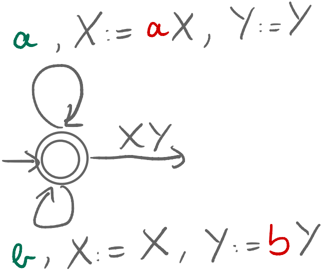
PROPERTIES

* effective translation between models
* decidable equivalence problem
* closed under composition

HOW TO COMPUTE A TRANSDUCTION?

* Sort: aababba ⟼ aaaabbb





GENERAL

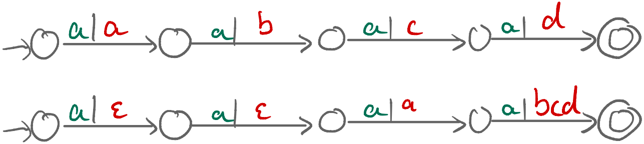
input ⟼ output : many ways to compute it!

⇒ REASON FOR MANY UNDECIDABILITY RESULTS

* equivalence undecidable for non-det. transducer (NFT)
* universality undecidable for NFTs
* ...

HOW TO COMPARE
TWO TRANSDUCER
COMPUTATIONS?

FINITE TRANSDUCERS: DELAY



Compare two computations c_1, c_2 on $aaaa$

$t=1 \quad d(a, \varepsilon) = 1$

$t=2 \quad d(ab, \varepsilon) = 2$

$t=3 \quad d(abc, a) = 2$

$t=4 \quad d(abcd, abcd) = 0$

$$\text{delay}(c_1, c_2) = 2$$

Delay between computations: How many symbols is one output ahead of the other?

FINITE TRANSDUCERS

ROBUSTNESS of DELAY

* functional NFTs T_1, T_2 : There is a computable K, s.t.
  if $T_1 \equiv T_2$, then $\text{delay}(T_1, T_2) \leq K$.

  (COMPLETENESS)

* Given K, the set

  $$\{ c_1 \otimes c_2 \mid \text{delay}(c_1, c_2) \leq K \} \text{ is regular.}$$

  computations

  (REGULARITY)

FINITE TRANSDUCERS

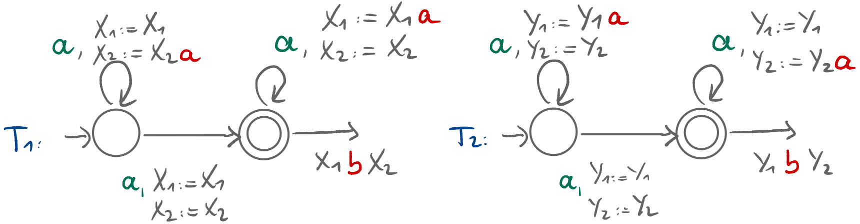
APPLICATIONS based on DELAY

* Decidable: Given k and some NFT T, is T k-valued?
  (i.e. for each input, there are at most k outputs)
* Decidable: Given some NFT T, is T finite-valued?
  [Weber '96] [Sakarovitch, de Souza '08]
* Decidable: Equivalence of finite-valued transducers
  [Filiot, Jecker, Löding, Winter '16]
* Decomposition Theorem: Every finite valued transducer is effectively equivalent to a finite union of 1-valued ones.
  [Weber '96] [Sakarovitch, de Souza '08]
* Canonical notions for input-deterministic & 1-valued transducers
  [Choffrut '03] [Reutenauer, Schützenberger '91]
* Characterization of sequential functions [Choffrut '77]
* Approximations of undecidable problems
  [Béal, Carton, Prieur, Sakarovitch '03]

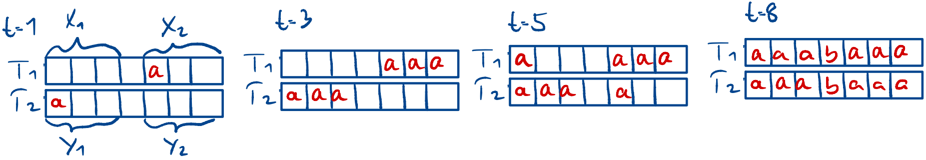
THE QUEST FOR A GOOD DELAY NOTION FOR STREAMING STRING TRANSDUCERS

A GOOD DELAY NOTION?

IDEA same notion as for finite transducers?

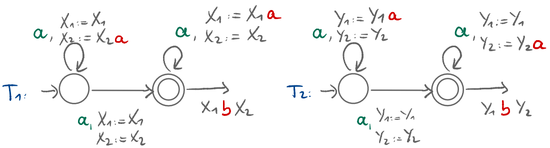


Compare computations on a^7, where each loop ×3



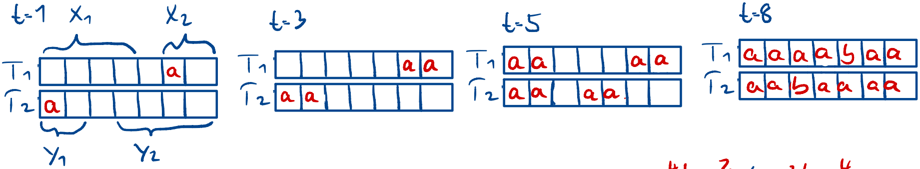
For $t=1,\dots,8$ outputs are the same (length).

A GOOD DELAY NOTION?



Compare on a^7

where $\ell_1 \times 2$, $\ell_2 \times 4$



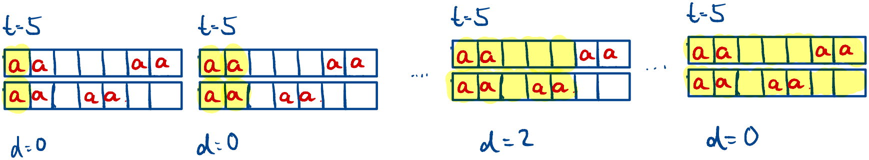
* For $t=1,..,7$ outputs are the same, for $t=8$: $a^4ba^2 \neq a^2ba^4$
* In general: final output only equal if $\#\ell_1 = \#\ell_2$
  - ⇝ not a regular property
  - ⇝ delay notion for NFTs does not reflect that

A GOOD DELAY NOTION?

TAKEAWAY

* SSTs do not build their output left-to-right
* Delay notion should reflect that

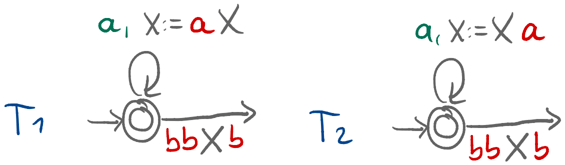
IDEA



* $d_{t=5} = \max\{0,0,\dots,2,\dots,0\}$
* $\text{delay} = \max_t d_t$

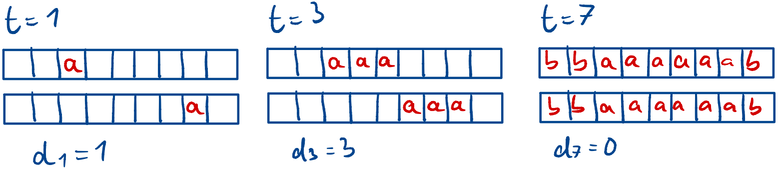
A GOOD DELAY NOTION?

Try the refined delay notion out!



$T_1 \equiv T_2$

$a^6 \mapsto bba^6b$



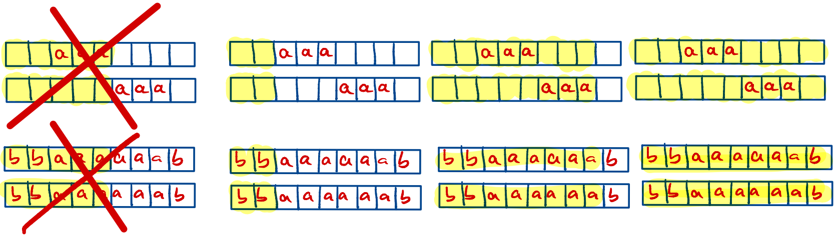
* $T_1 \equiv T_2$, but delay between computations grows with input length

A GOOD DELAY NOTION?

TAKEAWAY

* Periodic words: $u \cdot u^n = u^n \cdot u$
* Delay notion should reflect that!

=> Do not measure inside periodic factors:



DELAY BETWEEN SST COMPUTATIONS

DELAY NOTION FOR SST

DELAY NOTION

* sensitive to FINAL output placement
* immune against periodic factors

THEOREM

Filiot, Jecker, Löding, Winter

* Given SSTs T_1, T_2, there is a computable K such that if $T_1 \equiv T_2$, then $delay(T_1, T_2) \leq K$. (Completeness)

* Given K, the set $\{ c_1 \otimes c_2 \mid delay(c_1, c_2) \leq K \}$ is regular. (Regularity)

↗ SST computations

FURTHER RESULTS

THEOREM A complete and regular notion of delay for SSTs.

COROLLARY Delay notion is machine-independent.

Completeness result also holds for 2DFT, MSO-T, ...

THEOREM SST delay notion on finite transducers coincides with established delay notion.

APPLICATIONS

* Equivalence of NSST decidable up to fixed delay
* Register-minimization problem for NSST & SST decidable up to fixed delay.

SUMMARY

* Introduced a COMPLETE and REGULAR notion of delay for SSTs.

FUTURE WORK

* Decomposition theorem for SSTs via delay?
* Deciding finite-valuedness for SSTs via delay?