Output Strictly Local Functions

Jane Chandlee, Rémi Eyraud, Jeffrey Heinz

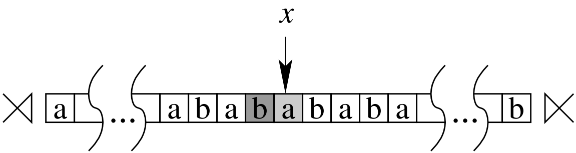
Haverford College | QARMA Team, LIF Marseille | University of Delaware

September 3, 2016

Objectives

- Define *Output Strictly Local (OSL) functions*, a subclass of subsequential string-to-string functions.
- Present a learning algorithm that learns any OSL function in quadratic time and data.
- Motivate the interest in this class with certain types of phonological input-output maps.

Strictly Local Stringsets



(McNaughton and Papert, 1971, Rogers and Pullum, 2011, Rogers et al., 2013)

Input Strictly Local functions

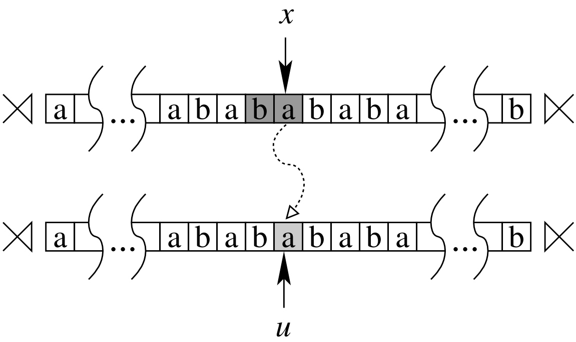


Figure: A function f is ISL-k if the output string u of each input element x depends only on x and the $k-1$ input element previous to x.

Myhill-Nerode theorem

- A stringset $\mathcal{L}$ is regular iff the tail equivalence relation partitions Σ^* into a finite set of equivalence classes.

$$\texttt{tails}_L(w) = \{u \mid wu \in L\}$$

Strictly Local stringsets

- $\mathcal{L}$ is Strictly Local iff $\forall w_1, w_2 \in \Sigma^*$, there exists $k \in \mathbb{N}$ such that if $\texttt{Suff}^{k-1}(w_1) = \texttt{Suff}^{k-1}(w_2)$ then $\texttt{tails}_L(w_1) = \texttt{tails}_L(w_2)$.

Input Strictly Local functions

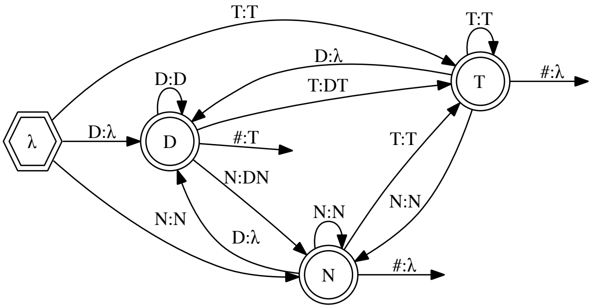
Definition (Tails w.r.t. f)

Let $f : \Sigma^* \to \Delta^*$ be a function. For $x \in \Sigma^*$, the tails of x, $\mathtt{tails}_f(x) = \{(y, v) \mid f(xy) = uv \wedge u = \mathtt{lcp}(f(x\Sigma^*))\}$

Definition (Input Strictly Local Functions)

A function $f : \Sigma^* \to \Delta^*$ is ISL if there is a k such that for all $u_1, u_2 \in \Sigma^*$, if $\mathtt{Suff}^{k-1}(u_1) = \mathtt{Suff}^{k-1}(u_2)$ then $\mathtt{tails}_f(u_1) = \mathtt{tails}_f(u_2)$.

Input Strictly Local functions



Final Devoicing: D# $\mapsto$ T#

Input Strictly Local functions

- The ISL functions are efficiently learnable from positive data.
  - Input Strictly Local Function Learning Algorithm (ISLFLA) (Chandlee et al., 2014) (quadratic time)
  - Structured Onward Subsequential Function Inference Algorithm (SOSFIA) (Jardine et al., 2014) (linear time)

Input Strictly Local functions

- Almost all phonological processes that apply locally can be modeled with ISL functions (Chandlee, 2014).
  - substitution
  - deletion
  - epenthesis
  - metathesis
  - local partial reduplication
- 95% of the approximately 5500 patterns in P-Base (Mielke, 2008, v1.95).

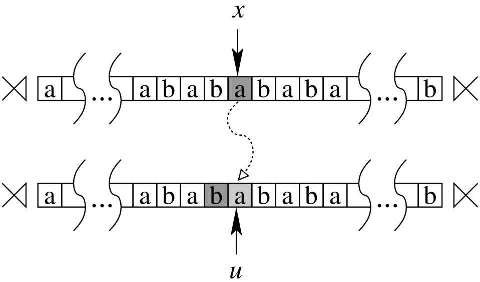
Input Strictly Local functions

- One notable class of exceptions is iterative spreading:

(1) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãw̃ãsan], ‘supervision’

(2) ISL version: /pəŋawasan/ ↦ *[pəŋãwasan]

Output Strictly Local Functions



Bigger picture

- Both ISL and OSL appear to be crucial components to the best characterization of locality in phonology.
- Restricting phonological maps computationally:
  1. enables efficient learning of phonological maps from positive data.
  2. provides insight into the kinds of transformations a theory of phonology should allow.

Prefix function

Definition (Prefix function)

Let $f : \Sigma^* \to \Delta^*$ be a subsequential function. Then the prefix function $f^p : \Sigma^* \to \Delta^*$ associated to f is $f^p(w) = \mathtt{lcp}(\{f(w\Sigma^*)\})$

$f(\text{AD}) = \text{AT}$
$f^p(\text{AD}) = \text{A}$

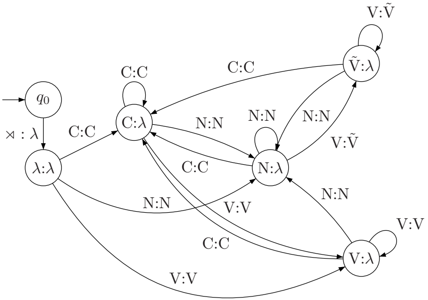
$f(\text{ADA}) = \text{ADA}$
$f^p(\text{ADA}) = \text{ADA}$

OSL Functions

Definition (Output Strictly Local Function)

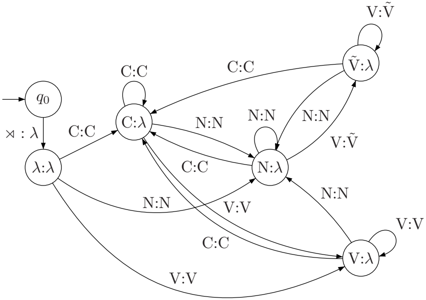
A subsequential function f is k-OSL if for all w_1, w_2 in Σ^*, $\mathtt{Suff}^{k-1}(f^p(w_1)) = \mathtt{Suff}^{k-1}(f^p(w_2)) \Rightarrow \mathtt{tails}_f(w_1) = \mathtt{tails}_f(w_2)$.

k-OSL transducers



/pəŋawasan/ ↦ [pəŋãw̃ãsan]

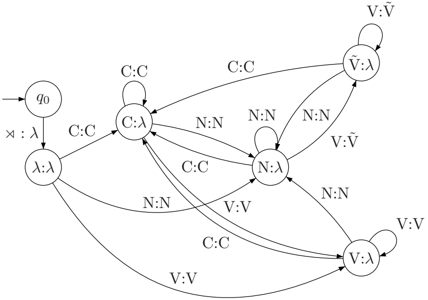
k-OSL transducers



/pəŋawasan/ ↦ [pəŋãw̃ãsan]

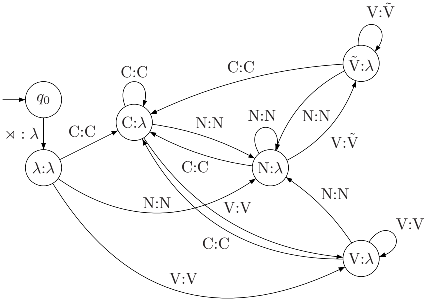
CVNVVVCVN ↦ CVNṼṼṼCVN

k-OSL transducers



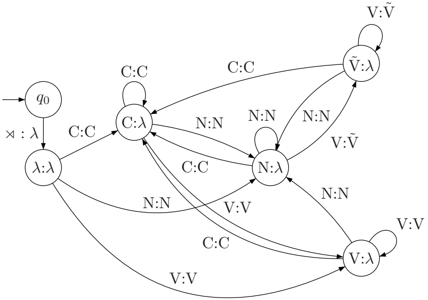
Delimited transducer: input strings augmented with start and end symbols (⋊*w*⋉), start state has single outgoing transition on ⋊, and states have an outgoing transition on ⋉to the final state.

k-OSL transducers



$$Q = \Delta^{\leq k-1} \cup \{q_0, q_f\}$$

k-OSL transducers



$$(\forall q \in Q, \forall a \in \Sigma, \forall u \in \Delta^*)\big[(q, a, u, q') \in \delta \implies q' = \texttt{Suff}^{k-1}(qu)\big]$$

k-OSL transducers

Lemma

Let T be an OSL-k DSFST. Then

$$(q_0, \rtimes w, u, q) \in \delta^* \iff f^p(w) = u$$

.

Learning definition

Definition (Characteristic sample)

For a $(\mathbb{T}, \mathbb{R})$-learning algorithm $\mathfrak{A}$, a sample CS is a *characteristic sample* of a function $t \in \mathbb{T}$ if for all samples S for t it is the case that $CS \subseteq S$ and $\mathfrak{A}$ returns a representation r such that $\mathcal{L}(r) = t$.

Learning definition

Definition (Identification in polynomial time and data)

A class $\mathbb{T}$ of functions is *identifiable in polynomial time and data* if there exists a $(\mathbb{T}, \mathbb{R})$-learning algorithm $\mathfrak{A}$ and two polynomials $p()$ and $q()$ such that:

1. For any sample S of size m for $t \in \mathbb{T}$, $\mathfrak{A}$ returns a hypothesis $r \in \mathbb{R}$ in $\mathcal{O}(p(m))$ time.
2. For each representation $r \in \mathbb{R}$ of size n, with $t = \mathcal{L}(r)$, there exists a characteristic sample of t for $\mathfrak{A}$ of size at most $\mathcal{O}(q(n))$.

Learning result

Theorem

OSLFIA identifies the k-OSL functions in quadratic time and data.

OSLFIA

Assumptions: Σ, Δ, and k are fixed

Input: set of string pairs $(\rtimes w \ltimes,\ v)$ consistent with the target function

Output: an OSL-k DSFST

OSLFIA

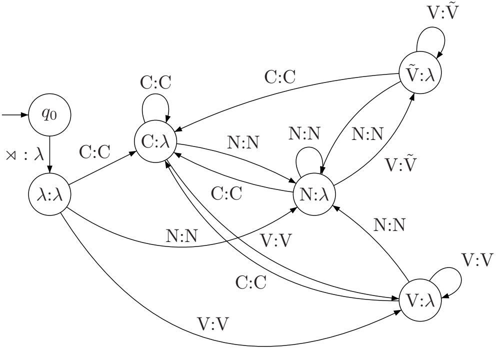
/pəŋawasan/ $\mapsto$ [pəŋãw̃ãsan]

$\Sigma = \{\text{C, V, N}\}$

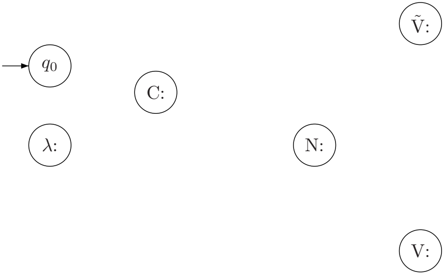
$\Delta = \{\text{C, V, }\tilde{\text{V}}\text{, N}\}$

$k = 2$

OSLFIA



OSLFIA



OSLFIA

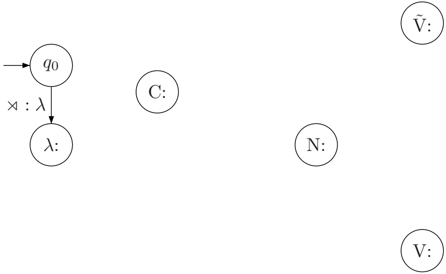
$S = \{\ (\rtimes C \ltimes, C),\ (\rtimes V \ltimes, V),\ (\rtimes N \ltimes, N),\ (\rtimes NV \ltimes, N\tilde{V}),\ \ldots$
$(\rtimes NVV \ltimes, N\tilde{V}\tilde{V}),\ \ldots\ \}$

OSLFIA

$S = \{$ ($\rtimes$C$\ltimes$,C), ($\rtimes$V$\ltimes$,V), ($\rtimes$N$\ltimes$,N), ($\rtimes$NV$\ltimes$,N$\tilde{V}$), ... ($\rtimes$NVV$\ltimes$,N$\tilde{V}\tilde{V}$), ... $\}$

Determine the *longest common prefix* (`lcp`) of all of the output strings in S.

OSLFIA

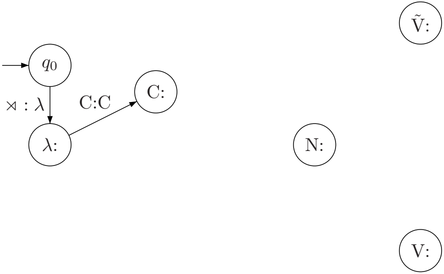


Destination state is the suffix of length $k - 1$ of the *output*.

OSLFIA

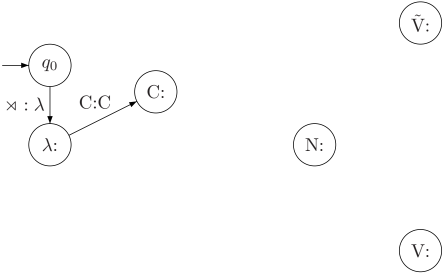
- From this destination state, identify the shortest input ‘extensions’ (i.e., C, N, V, etc.).
- For each input extension, identify the `lcp` of the output strings for the inputs in the dataset that begin with that prefix.

OSLFIA



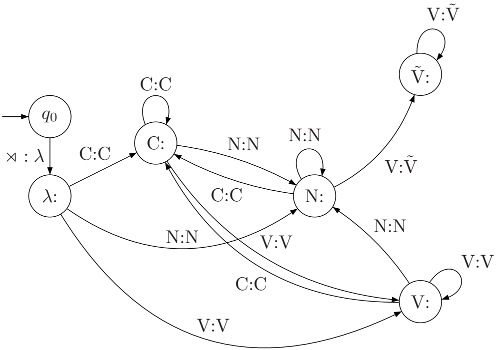
$S = \{$ ... (C,C), (CV,CV), (CN,CN), ...$\}$

OSLFIA



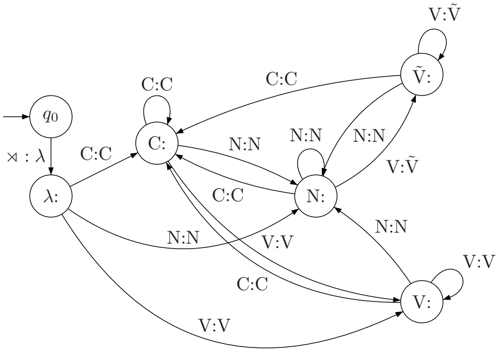
$S = \{ \ldots \ (C,C),\ (CV,CV),\ (CN,CN),\ \ldots \}$

OSLFIA

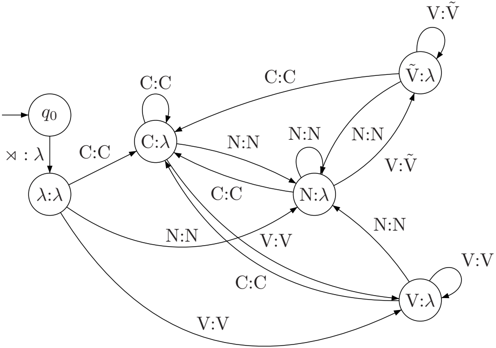


$S = \{ \ldots$ (NV,NṼ), (NVV,NṼṼ), (NVVV,NṼṼṼ), $\ldots\}$

OSLFIA



OSLFIA


q_0
⋊ : λ
λ:λ
C:C
C:λ
C:C
N:N
N:λ
N:N
C:C
$\tilde{V}$:λ
V:$\tilde{V}$
C:C
N:N
V:$\tilde{V}$
N:N
V:V
C:C
V:λ
V:V
N:N
V:V

Learning result

Theorem

OSLFIA identifies the k-OSL functions in quadratic time and data.

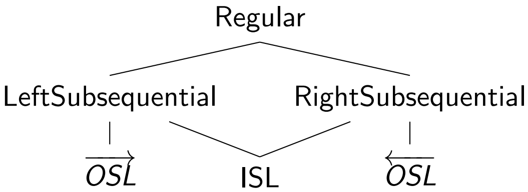
- More efficient than OSTIA, which learns total subsequential functions in cubic time (Oncina et al., 1993) and OSTIA-R, which learns partial subsequential functions (Castellanos et al., 1998).

Back to Phonology

- Johore Malay nasal spreading is *progressive* (i.e., left-to-right).
- Regressive (right-to-left) spreading
  (Ajíbóyè, 2001, Ajíbóyè and Pulleyblank, 2008, Walker, 2014):

  (3) Mòbà Yoruba
  /ujĩ/ $\mapsto$ [ũj̃ĩ], 'praise(n.)'

Back to Phonology



What's left?

- In rule-based terms, ISL functions can model rules with two-sided contexts that apply *simultaneously*.
- OSL functions can model rules with one-sided contexts that apply *left-to-right* or *right-to-left*.

(4) $\quad$ A $\rightarrow$ B / C __ (left-to-right)

(5) $\quad$ A $\rightarrow$ B / __ D (right-to-left)

Two-sided iterative rules

(6) French Schwa Deletion

ə → ∅ / VC __ CV

(7) a. /ty dəvənɛ/ ↦ [ty dvənɛ], 'you became'

b. /ty dəvənɛ/ ↦ [ty dəvnɛ]

c. /ty dəvənɛ/ ↦ *[ty dvnɛ]

(Kaplan and Kay, 1994, Dell, 1985, Dell, 1980, Dell, 1973, Noske, 1993)

Long-distance processes

(8) Kikongo (Rose and Walker, 2004)
/tu+nik+idi/ $\mapsto$ [tunikini] 'we ground'

Current and Future Work

- Iterative rules with two-sided contexts can be modeled with **Input-Output SL functions**, in which one side is matched to the input and the other to the output.
- Long-distance processes may be describable with additional classes of subregular relations based on language classes other than SL (i.e., SP)
  (Heinz, 2010, Rogers et al., 2010, Rogers et al., 2013).

Conclusions

- The OSL functions are a proper subset of the subsequential functions based on the notion of locality that defines the SL stringsets.
- These functions correspond exactly to a particular type of transducers, the OSL DSFSTs.
- The OSLFIA exploits the structure of this class to efficiently learn any OSL-k mapping.

Thank you!

References I

Ajíbóyè, O. (2001).
Nasalization in Mòbà.
In Oh, S., Sawai, N., Shiobara, K., and Wojdak, R., editors, *Proceedings of the Northwest Linguistics Conference*, pages 1–18. University of British Columbia Working Papers in Linguistics 8. Vancouver: University of British Columbia, Department of Linguistics.

Ajíbóyè, O. and Pulleyblank, D. (2008).
Mòbà nasal harmony.
Ms., University of Lagos and University of British Columbia.

References II

Castellanos, A., Vidal, E., Varó, M. A., and Oncina, J. (1998).
Language understanding and subsequential transducer learning.
Computer Speech and Language, 12:193–228.

Chandlee, J. (2014).
Strictly Local Phonological Processes.
PhD thesis, The University of Delaware.

Chandlee, J., Eyraud, R., and Heinz, J. (2014).
Learning strictly local subsequential functions.
Transactions of the Association for Computational Linguistics, 2:491–503.

References III

Dell, F. (1973).
Les régles et les sons.
Paris: Hermann.

Dell, F. (1980).
Generative phonology and French phonology.
Cambridge: Cambridge University Press.

Dell, F. (1985).
Les régles et les sons.
Paris: Hermann, 2 edition.

Heinz, J. (2010).
Learning long-distance phonotactics.
Linguistic Inquiry, 41(4):623–661.

References IV

Jardine, A., Chandlee, J., Eyraud, R., and Heinz, J. (2014).
Very efficient learning of structured classes of subsequential functions from positive data.
In Clark, A., Kanazawa, M., and Yoshinaka, R., editors, *Proceedings of the Twelfth International Conference on Grammatical Inference (ICGI 2014)*, volume 34, pages 94–108. JMLR: Workshop and Conference Proceedings.

Kaplan, R. and Kay, M. (1994).
Regular models of phonological rule systems.
Computational Linguistics, 20(3):331–378.

References V

McNaughton, R. and Papert, S. (1971).
Counter-Free Automata.
MIT Press.

Mielke, J. (2008).
The Emergence of Distinctive Features.
Oxford University Press, Oxford.

Noske, R. (1993).
A theory of syllabification and segmental alternation.
Niemeyer, Tübingen.

References VI

Oncina, J., García, P., and Vidal, E. (1993).
Learning subsequential transducers for pattern recognition tasks.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 15:448–458.

Onn, F. M. (1980).
Aspects of Malay Phonology and Morphology: A Generative Approach.
Kuala Lumpur: Universiti Kebangsaan Malaysia.

References VII

Rogers, J., Heinz, J., Bailey, G., Edlefsen, M., Visscher, M., Wellcome, D., and Wibel, S. (2010).
On languages piecewise testable in the strict sense.
In Ebert, C., Jäger, G., and Michaelis, J., editors, *The Mathematics of Language*, volume 6149 of *Lecture Notes in Artifical Intelligence*, pages 255–265. Springer.

Rogers, J., Heinz, J., Fero, M., Hurst, J., Lambert, D., and Wibel, S. (2013).
Cognitive and sub-regular complexity.
In Morrill, G. and Nederhof, M.-J., editors, *Formal Grammar*, volume 8036 of *Lecture Notes in Computer Science*, pages 90–108. Springer.

References VIII

Rogers, J. and Pullum, G. (2011).
Aural pattern recognition experiments and the subregular hierarchy.
Journal of Logic, Language and Information, 20:329–342.

Rose, S. and Walker, R. (2004).
A typology of consonant agreement as correspondence.
Language, 80:475–531.

Walker, R. (2014).
Nonlocal trigger-target relations.
Linguistic Inquiry, 45(3):501–523.