Learning Opaque Maps

Jane Chandlee, Jeffrey Heinz, Adam Jardine

janemc@udel.edu heinz@udel.edu ajardine@udel.edu

Central Contribution

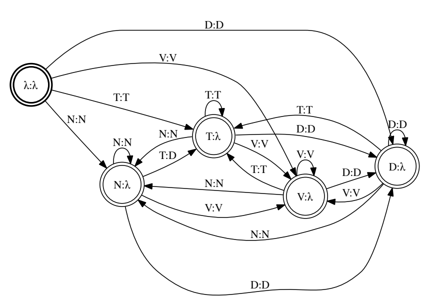
Representing (and therefore learning) opaque maps is difficult in classic Optimality Theory (Prince and Smolensky 1993/2004) and many of its variants. **We present two learning algorithms that learn the opaque maps that Baković (2007) presents to illustrate the typology of opaque phenomena in phonology**. These algorithms are provably correct in a well-defined sense. They succeed because the target maps are Input Strictly Local (ISL), which is a computational property of maps known to hold for many phonological processes (Chandlee, 2014). Here it is shown that the ISL property also holds for opaque maps, which are not normally thought of as corresponding to singular processes.

Computational Approach

- Both rule- and constraint-based theories of generative phonology concur on the existence of a map from input (underlying) to output (surface) forms.
- We model these maps with *functions* with the goal of identifying computational properties that are independent of grammatical formalisms (such as rules and constraints).

  (1) Let f be a function which voices obstruents after nasals. Example: f(kampa) = [kamba]

  Q: What class of functions does f belong to?
- Identifying the most *restrictive* set of functions needed for phonological maps leads to a better characterization of the components of phonological grammars.
- The restrictive nature of the classes may also enable correct and efficient learning from positive data (i.e., the properties of the class structure the learner's hypothesis space) (Heinz, 2009, 2010).
- FSTs are a finite means of representing an infinite function like (1) (Johnson, 1972; Kaplan and Kay, 1994; Mohri, 1997).

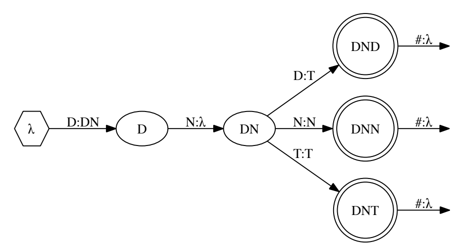


	k		a		m		p		a		#	
λ	$\Rightarrow$	T	$\Rightarrow$	V	$\Rightarrow$	N	$\Rightarrow$	T	$\Rightarrow$	V	$\Rightarrow$	
	k		a		m		b		a		λ	

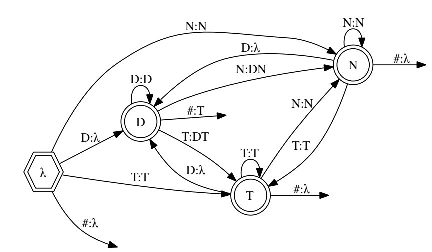
- Particular classes of functions correspond to particular types of FSTs.
- Phonological maps can be modeled with the FSTs that correspond to the *Input Strictly Local* functions, and this class is efficiently learnable (Chandlee, 2014; Chandlee et al., 2014).
- Informally, a function is Input Strictly k-Local if the output of every input string $a_0a_1\cdots a_n$ is $u_0u_1\cdots u_n$ where u_i is a string which only depends on a_i and the $k-1$ symbols before a_i in the input string (so $a_{i-k+1}a_{i-k+2}\cdots a_{i-1}$).
- ISL $\subset$ Subsequential $\subset$ Regular (Mohri, 1997; Chandlee, 2014; Chandlee et al., 2014).
- We show that opaque maps also belong to the ISL class and are therefore also learnable.

The Algorithms

1. **Input Strictly Local Function Learning Algorithm (ISLFLA)**: state merging approach (Chandlee, 2014; Chandlee et al., 2014)
   - Generalizes from input data by merging states, and the state merging criterion is based on the defining property of ISL functions (merge states with the same $k-1$ suffix).

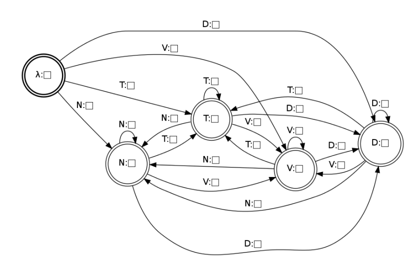


PTT for final devoicing

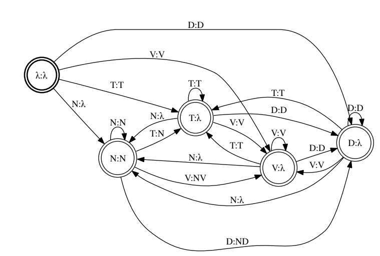


Final devoicing ISL-FST

2. **Structured Onward Subsequential Function Inference Algorithm (SOSFIA)**: 'output-empty transducer' approach (Jardine et al., 2014)
   - The eFST defines the *class of functions* in the range of the learner, or (in phonological terms) the eFST is the *range of constraints* for which it is possible to learn repairs.



eFST for $\Sigma = \{D,V,N,T\}$ and $k=2$



Indonesian fusion (NT $\mapsto$ N)

Opaque maps

Baković (2007) discusses 6 types of phonological opacity.

- Counterbleeding (McCarthy, 1999)

  (2) Yokuts

	'might fan'
	/ʔili:+l/
+long → -high	ʔile:l
V → -long / __ C #	ʔilel
	[ʔilel]

- Counterfeeding-on-environment (McCarthy, 1999)

  (3) Bedouin Arabic

	'Bedouin'
	/badw/
a → i / __ σ	badw
G → V / C __#	badu
	[badu]

- Counterfeeding-on-focus (McCarthy, 1999)

  (4) Bedouin Arabic

	'he wrote'
	/katab/
i → ∅ / __ σ	katab
a → i / __ σ	kitab
	[kitab]

- Self-destructive feeding

  (5) Turkish (Sprouse, 1997)

	'your baby'
	/bebek+n/
∅ → i / C __ C #	bebekin
k → ∅ /V __ +V	bebein
	[bebein]

Opaque maps (cont.)

- Non-gratuitous feeding

  (6) Classical Arabic (McCarthy, 2007)

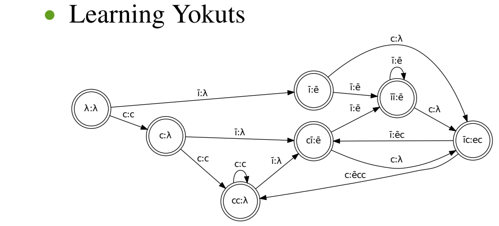
	'write.MS'
	/ktub/
$\emptyset \rightarrow V_i$ / # __ CCV_i	uktub
∅ → ʔ / # __ V	ʔuktub
	[ʔuktub]

- Cross-derivational feeding

  (7) Lithuania (Baković, 2005; Odden, 2005)

	'to strew all over'
	/ap-berti/
∅ → i / C __ C	apiberti
-voice → +voice / __ D	apiberti
	[apiberti]

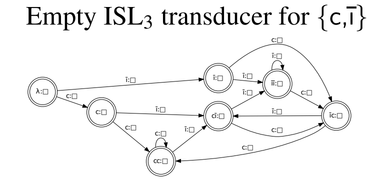
- Learning Yokuts



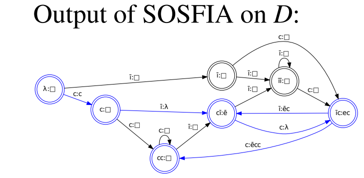
Sample data D

cīcī	:	cēcē
cīc	:	cec
cīcc	:	cēcc
cī	:	cē

Empty ISL₃ transducer for {c,ī}



Output of SOSFIA on D:



Discussion

- When viewed as single maps, all of these cases of opacity are ISL for some k and therefore fall within the class of functions learnable by these algorithms.
- Like OT and other constraint-based formalisms, this approach does not factor the phonological grammar into individual processes.
- Unlike these constraint-based theories, this approach can handle many different types of opacity.

Remaining issues

- Learning URs: both algorithms must be given data in the form of a set of (UR, SR) pairs. How can URs instead be learned (cf. Tesar (2014))?
- Long-distance maps are not ISL and therefore either 1) must be decomposed into two ISL maps or 2) are restricted by a distinct but comparable computational property (see Heinz, 2009, 2010).

Selected References

Baković, E. (2005). Antigemination, assimilation and the determination of identity. *Phonology*, 22:279–315.
Baković, E. (2007). A revised typology of opaque generalisations. *Phonology*, 24(2):217–259.
Chandlee, J. (2014). *Strictly Local Phonological Processes*. PhD thesis, U. of Delaware.
Chandlee, J., Eyraud, R., and Heinz, J. (2014). Learning strictly local subsequential transducers. *TACL 2*, pages 491–503.
Heinz, J. (2009). On the role of locality in learning stress patterns. *Phonology*, 26:303–351.
Heinz, J. (2010). Learning long-distance phonotactics. *Linguistic Inquiry*, 41:623–661.
Jardine, A., Chandlee, J., Eryaud, R., and Heinz, J. (2014). Very efficient learning of structured classes of subsequential functions from positive data. In *Proceedings of the 12th International Conference on Grammatical Inference (ICGI 2014)*, JMLR Workshop Proceedings, pages 94–108.
Johnson, C. D. (1972). *Formal aspects of phonological description*. Mouton.
Kaplan, R. and Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics*, 20:331–78.
McCarthy, J. J. (1999). Sympathy and phonological opacity. *Phonology*, 16:331–399.
McCarthy, J. J. (2007). *Hidden generalizations : phonological opacity in Optimality Theory*. London: Equinox.
Mohri, M. (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, 23(2):269–311.
Odden, D. (2005). *Introducing phonology*. Cambridge: Cambridge University Press.
Sprouse, R. (1997). A case for enriched inputs. *Available as ROA-193 from the Rutgers Optimality Archive*.
Tesar, B. (2014). *Output-driven Phonology*. Cambridge University Press.