Learning Repairs for Marked Structures

Jane Chandlee, Adam Jardine, and Jeffrey Heinz
chandlee@asel.udel.edu ajardine@udel.edu heinz@udel.edu

Central Contribution

We present a theory of phonology based on the computational properties of input-output mappings. These properties define restrictive classes of mappings which are learnable by a provably correct, efficient algorithm. The algorithm learns both the active surface constraints and the repairs, including opaque mappings.

Introduction

- Output-oriented theories of phonological grammars are in part motivated by 'conspiracies' in which the same marked structure is targeted by a range of different processes.

(1) Fusion (Indonesian)
/məN+pilih/ ↦ [məmilih], 'to choose'

(2) Voicing (Quechua)
/kam+pa/ ↦ [kamba], 'yours'

(3) Denasalization (Toba Batak) (Hayes, 1986)
/maŋinum tuak/ ↦ [maŋinup tuak], 'drink palm wine'

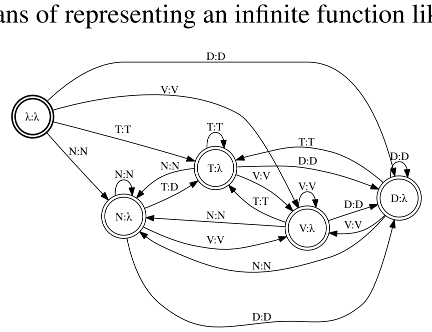
- In OT a single markedness constraint (*NC̥) ranked with respect to a set of faithfulness contraints accounts for this variation (Pater, 2004).
- We likewise propose a learner that separates the marked structure and the repair and can furthermore learn opaque input-output mappings.

Computational Approach

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

(4) $f(\text{kam+pa}) = [\text{kamba}]$

- In particular, identifying the most *restrictive* computational properties leads us to a better characterization of the components of phonological grammars (Johnson, 1972; Kaplan and Kay, 1994; Mohri, 1997).
- FSTs are a finite means of representing an infinite function like (4).



	k		a		m		p		a		#
λ	⇒	T	⇒	V	⇒	N	⇒	T	⇒	V	⇒
	k		a		m		b		a		λ

Figure 1: Quechua obstruent voicing

- This FST belongs to a restricted class called *Input Strictly Local* (ISL ⊂ Subsequential ⊂ Regular) (Mohri, 1997; Chandlee, 2014; Chandlee et al., 2014).

The Algorithm

- **Structured Onward Subsequential Function Inference Algorithm (SOSFIA)** (Jardine et al., 2014)
- Input: set of input/output pairs (data) and **an output-empty subsequential FST (eFST)** (structure)

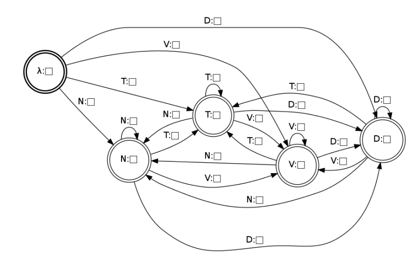
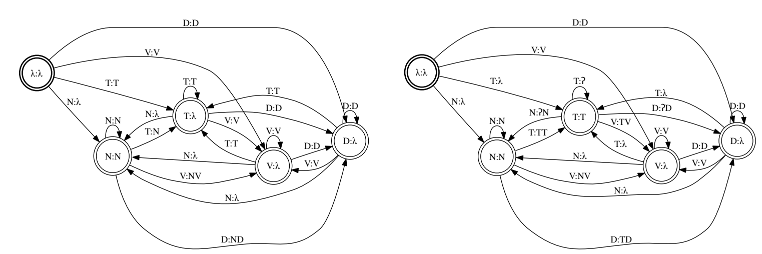


Figure 2: An output-empty FST for learning repairs with T, D, V, and N

- SOSFIA uses data input/output pairs to fill in the *blank outputs* of the eFST.
- The eFST defines the *class of functions* in the range of the learner.



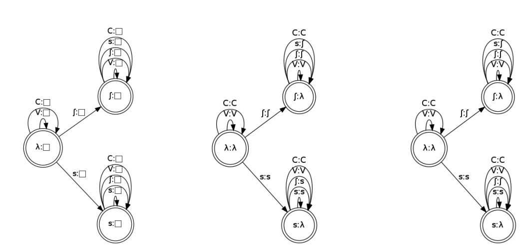
Indonesian fusion   Toba Batak denasalization

Figure 3: Two FSTs in range of SOSFIA, given Figure 2

- In a phonological context, this determines the *range of constraints* for which it is possible to learn repairs—with Figure 2, SOSFIA only learns *local* constraints concerning the natural classes of vowels, nasals, and voiced and voiceless obstruents.

Advantages

- Provably learns any transduction in the class defined by eFST from positive data.
- Like ISLFLA of Chandlee (2014); Chandlee et al. (2014), but uses knowledge of structure to be very efficient: the learner is fast (linear in size of input) and requires a small amount of data (linear in size of eFST).
- Consistent with research showing phonological processes to be subsequential (Chandlee and Heinz, 2012; Heinz and Lai, 2013; Payne, 2014), with the majority in the more restrictive ISL class (Chandlee. 2014).
- Learns opaque processes without recourse to intermediate representations (see Figure 3).
- Can learn long-distance phenomena if supplied the correct eFST



eFST   Samala (Applegate, 1972)   Sarcee (Cook, 1978)

Figure 4: FSTs for long-distance sibilant harmony typology

Discussion

- The eFST represents the learner's *a priori* knowledge about the target mapping it is trying to learn.
- When the target is part of a highly-structured class of mappings (such as the ISL ones), this *a priori* knowledge greatly facilitates learning.
- In the case of phonology, a process is ISL provided it can be described with a rule that applies simultaneously in which the target and triggering context are a contiguous substring of bounded length (Chandlee, 2014; Chandlee and Heinz, 2014).
- Like classical OT,
  1) there are no intermediate representations,
  2) the constraints are given *a priori* (here in the structure of the class), and
  3) we can provide strong learnability results.
- Unlike classical OT,
  1) opaque mappings are representable and learnable,
  2) non-regular mappings cannot be generated
  (cf. Riggle, 2004; Gerdemann and Hulden, 2012).

Future and Ongoing Work

- One of these classes, the ISL class, is well defined (Chandlee, 2014; Chandlee et al., 2014). Non-local classes (such as in Figure 4) are not well understood. What non-local class(es) is (are) there? How do they relate to the ISL class?
- Given that there are multiple classes towards which human learners are biased, they learn different kinds of mappings. How are these mappings composed into a single, unified phonology? This is a different question than the 'rule-ordering' composition of, ex., Kaplan and Kay (1994).
- Given that optimization of simple constraints under particular rankings can result in complex mappings (Riggle, 2004; Gerdemann and Hulden, 2012), it appears unlikely that OT can account for the computational properties of phonological mappings mentioned here, but this certainly deserves further investigation.

References

Applegate, R. (1972). *Inseño Chumash grammar*. PhD thesis, University of California, Berkeley.

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. *Transactions of the Association for Computational Linguistics*, page to appear.

Chandlee, J. and Heinz, J. (2012). Bounded copying is subsequential: Implications for metathesis and reduplication. In *Proceedings of the 12th Meeting of the ACL Special Interest Group on Computational Morphology and Phonology*, pages 42–51, Montreal, Canada. Association for Computational Linguistics.

Chandlee, J. and Heinz, J. (2014). Strictly local phonological processes. under review.

Cook, E.-D. (1978). Palatalizations and related rules in Sarcee. In Cook, E.-D. and Kaye, J., editors, *Linguistic studies of native Canada*, pages 19–35. Vancouver: University of British Columbia Press.

Gerdemann, D. and Hulden, M. (2012). Practical finite state optimality theory. In *Proceedings of the 10th International Workshop on FSMNLP*, pages 10–19. ACL.

Hayes, B. (1986). Assimilation as spreading in Toba Batak. *LI*, 17(3):467–499.

Heinz, J. and Lai, R. (2013). Vowel harmony and subsequentiality. In Kornai, A. and Kuhlmann, M., editors, *Proceedings of the 13th Meeting on Mathematics of Language*, Sofia, Bulgaria.

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.

Mohri, M. (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, 23(2):269–311.

Pater, J. (2004). Austronesian nasal substitution and other NC effects. In McCarthy, J., editor, *Optimality Theory in Phonology: A Reader*, pages 271–289. Oxford and Malden, MA: Blackwell.

Payne, A. (2014). Dissimilation as a subsequential process. In *Proceedings of the 44th Meeting of the North East Linguistic Society (NELS)*.

Riggle, J. (2004). *Generation, Recognition, and Learning in Finite State Optimality Theory*. PhD thesis, UCLA.